

GenCore version 5.1.3
Copyright (c) 1993 - 2002 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: December 9, 2002, 16:36:47 ; Search time 13952 Seconds

(without alignments)
8925.667 Million cell updates/sec

Title: US-09-672-725C-1

Perfect score: 4279
Sequence: 1 ggaagcgcagctcggaatg.....caaaaaaaaaaaaaa 4279Scoring table: IDENTITY_NUC
Gapop 10.0 , Gapext 1.0

Searched: 2054640 seqs, 14551402878 residues

Total number of hits satisfying chosen parameters: 4109280

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Database :

Listing first 45 summaries

GenBank1:*

- 1: gb_ba:*
- 2: gb_hlg:*
- 3: gb_in:*
- 4: gb_om:*
- 5: gb_ov:*
- 6: gb_pat:*
- 7: gb_ph:*
- 8: gb_pl:*
- 9: gb_pr:*
- 10: gb_ro:*
- 11: gb_sts:*
- 12: gb_sy:*
- 13: gb_un:*
- 14: gb_vt:*
- 15: em_ba:*
- 16: em_fun:*
- 17: em_hum:*
- 18: em_in:*
- 19: em_mu:*
- 20: em_om:*
- 21: em_or:*
- 22: em_ov:*
- 23: em_pat:*
- 24: em_ph:*
- 25: em_pl:*
- 26: em_ro:*
- 27: em_sts:*
- 28: em_un:*
- 29: em_vt:*
- 30: em_hlg_hum:*
- 31: em_hlg_inv:*
- 32: em_hlg_other:*
- 33: em_hlg_mus:*
- 34: em_hlg_pln:*
- 35: em_hlg_rtd:*
- 36: em_hlg_mam:*
- 37: em_hlg_vtc:*
- 38: em_sy:*
- 39: em_hlgo_hum:*
- 40: em_hlgo_mus:*
- 41: em_hlgo_other:*

score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	4279	100.0	4279	6 AX105057	AX105057 Sequence
2	4277.4	100.0	4279	6 AX105078	AX105078 Sequence
3	4275.8	99.9	4279	6 AX105080	AX105080 Sequence
4	4272.6	99.9	4279	6 AX105082	AX105082 Sequence
5	4221.8	98.7	4317	4 AF045016	AF045016 Canis fam
6	4221.8	98.7	4317	6 AX105059	AX105059 Sequence
7	3847.4	89.4	4045	12 AF269224	AF269224 Synthetic
8	3824	89.4	3934	4 CEFA19568	AF419568 Canis fam
9	3342.8	78.1	4669	6 E02326	E02326 Sequence 3
10	3341.2	78.1	4378	6 E02326	E02326 Multidrug r
11	3333.2	77.9	4669	6 AR091275	AR091275 Sequence
12	3333.2	77.9	4669	6 AR203322	AR203322 Sequence
13	3327.6	77.8	4646	6 AX391089	AX391089 Sequence
14	3324.4	77.7	4646	6 AX336420	AX336420 Sequence
15	3324.4	77.7	4646	6 AX336708	AX336708 Sequence
16	3324.4	77.7	4646	6 I49610	I49610 Sequence 2
17	3324.4	77.7	4646	6 HUMMDR1	M14758 Homo sapien
18	3315.6	77.5	4669	6 AR055785	AR055785 Sequence
19	3311.8	77.4	4186	6 AX108654	AX108654 Sequence
20	3311.2	77.4	4195	6 AX108656	AX108656 Sequence
21	3270.6	76.4	4264	6 AR051647	AR051647 Sequence
22	3270.6	76.4	4264	6 AR051650	AR051650 Sequence
23	3265.6	75.3	4192	9 AF016535	AF016535 Homo sapl
24	3247.6	75.9	3988	6 AX024454	AX024454 Sequence
25	3239.6	75.7	6505	6 AR028671	AR028671 Sequence
26	3239.6	75.7	9318	6 AR028672	AR028672 Sequence
27	3226.6	75.4	8630	6 AX012321	AX012321 Sequence
28	3223.4	75.3	8630	6 AX012320	AX012320 Sequence
29	3202	74.8	3860	6 AX322787	AX322787 Sequence
30	3200.4	74.8	3860	6 AX322789	AX322789 Sequence
31	3190.2	74.6	3489	6 AX481416	AX481416 Sequence
32	3056.8	71.4	3489	4 AB029153	AB029153 Felis cat
33	3001	70.1	3858	4 OA078609	OA078609 Ovis aries
34	2989.6	69.9	4296	10 CRUPGP1	M60040 C.grienseu p
35	2988	68.9	4304	10 CRUPGP1165	M59253 Chinese ham
36	2949	68.8	4927	10 AF257746	AF257746 Rattus no
37	2946	68.8	4356	10 MUSMDRAA	M30597 Mouse multi
38	2943.4	68.8	4924	10 MUSMDRIA	M33581 Mouse multi
39	2942.2	68.8	4323	10 AF286167	AF286167 Rattus no
40	2936.2	68.6	4788	6 AX322793	AX322793 Sequence
41	2846.8	66.5	3987	10 CRUPGP1185	M59254 Chinese ham
42	2787.4	65.1	4280	10 CRUPGP11	M60041 C.grienseu p
43	2724.6	63.7	4298	10 MUSMDR	M14757 Mouse multi
44	2716.6	63.5	4189	6 AX322791	AX322791 Sequence
45	2705	63.2	4233	6 AR123273	AR123273 Sequence

ALIGNMENTS

RESULT 1
AX105057
LOCUS AX105057 4279 bp
DEFINITION Sequence 1 from Patent WO0123540.
ACCESSION AX105057
VERSION AX105057.1 GI:13921209
KEYWORDS
SOURCE dog;
ORGANISM Canis familiaris
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Carnivora; Fissipedia; Canidae; Canis.
REFERENCE
AUTHORS 1 (bases 1 to 4279)
Stocker,P.J., Stetmel-Crespi,D.T., Crespi,C.L., Relf,T.C. and
Pattern,C.J.
TITLE P-glycoproteins and uses thereof

Pred. No. is the number of results predicted by chance to have a

Query Match	100.0%	Score 4277.4	DB 6	Length 4279
Best Local Similarity	100.0%	Prod. No. 0		
Matches 4278	Conservative 0	Mismatches 1	Indels 0	Gaps 0
QY	1	GGAGCGCGAGGTCGGGTGGATCTCGAAGGAGCGCCGTAAGGGAGTGGAGAGAACTT	60	
Db	1	GGAGCGCGAGGTCGGGTGGATCTCGAAGGAGCGCCGTAAGGGAGTGGAGAGAACTT	60	
QY	61	CTGGAAAAATGGGCAAAAAAGTAAATAATGAGACAGAAAGAAACCACTGTGAG	120	
Db	61	CTGGAAAAATGGGCAAAAAAGTAAATAATGAGACAGAAAGAAACCACTGTGAG	120	
QY	121	CACGTTTCGCAATGTTTGCATTCAAATTTGGCTCGATAGGTTGTATATGTTGTGGGGAC	180	
Db	121	CACGTTTCGCAATGTTTGCATTCAAATTTGGCTCGATAGGTTGTATATGTTGTGGGGAC	180	
QY	181	AATGGCTGCACATTCGCATGAGGCTGCACCTCCCTCATGATAGTCGCTTTTGGAAACAT	240	
Db	181	AATGGCTGCACATTCGCATGAGGCTGCACCTCCCTCATGATAGTCGCTTTTGGAAACAT	240	
QY	241	GACAGATGCTTTCGCAATGAGAGGAAATTTCAAGAAACAAACCTTTCCAGTTAATATTA	300	
Db	241	GACAGATGCTTTCGCAATGAGAGGAAATTTCAAGAAACAAACCTTTCCAGTTAATATTA	300	
QY	301	TGAAAGTTATCGAACAATACCAACTTTCATACCATCTGGAGAGGAAATGACAC	360	
Db	301	TGAAAGTTATCGAACAATACCAACTTTCATACCATCTGGAGAGGAAATGACAC	360	
QY	361	GTAATGCTATTATTATACAGTGGAGTGGTGGCTGGGCTGGTGGCTGTCATCAGAT	420	
Db	361	GTAATGCTATTATTATACAGTGGAGTGGTGGCTGGGCTGGTGGCTGTCATCAGAT	420	
QY	421	TTTCATTTCTGGTGGCTGGAGAGGAGACATATCTCAAAATTAATAAACAATTTTTTCA	480	
Db	421	TTTCATTTCTGGTGGCTGGAGAGGAGACATATCTCAAAATTAATAAACAATTTTTTCA	480	
QY	481	TGCTATCATGAGAGAGATGGCTGGTTTGAAGTGCATGACGCTGGGGAGCTTAAC	540	
Db	481	TGCTATCATGAGAGAGATGGCTGGTTTGAAGTGCATGACGCTGGGGAGCTTAAC	540	
QY	541	CCGGCTCACAGACATGCTCCAAAATTAATGAAGAAATTTGGCGCAAAATTTGGAAATGT	600	
Db	541	CCGGCTCACAGACATGCTCCAAAATTAATGAAGAAATTTGGCGCAAAATTTGGAAATGT	600	
QY	601	CTTTGACTCATATAGCAACATTTTTACCGGTTTATATAGTGGGTTTACACGTGGTGGAA	660	
Db	601	CTTTGACTCATATAGCAACATTTTTACCGGTTTATATAGTGGGTTTACACGTGGTGGAA	660	
QY	661	GGTAACTCTGTGATTTTGGCCATCAGCCCTGTCTTGAAGCTTTCAAGCGGCATCTGGGC	720	
Db	661	GGTAACTCTGTGATTTTGGCCATCAGCCCTGTCTTGAAGCTTTCAAGCGGCATCTGGGC	720	
QY	721	AAAGTACTACTCTCATATAGATAAAGAACTTGGCTATGCAAAAGCTGGAGCACT	780	
Db	721	AAAGTACTACTCTCATATAGATAAAGAACTTGGCTATGCAAAAGCTGGAGCACT	780	

Db	2941	TGTTCTTTTGGTAATCTCAGCTATTGTCCTTTGGTCCATGCGAGTGGGCGAGTCAAGTTC	3000
Qy	3001	ATTTCGCTCGACTATAGCCAAAGCCAAAGATATCAGCAGCCAGCTCATCATATTTGA	3060
Db	3001	ATTTCGCTCGACTATAGCCAAAGCCAAAGATATCAGCAGCCAGCTCATCATATTTGA	3060
Qy	3061	AAAAAGCCTCTGATTGACAGCTACAGGCCCTCAGGGCTCAAGCCAAATACGTTGAAGG	3120
Db	3061	AAAAAGCCTCTGATTGACAGCTACAGGCCCTCAGGGCTCAAGCCAAATACGTTGAAGG	3120
Qy	3121	AAATGTACATTTAATAGAGCTGTTCACATATCCACTCGACACAGACATCCCGTCT	3180
Db	3121	AAATGTACATTTAATAGAGCTGTTCACATATCCACTCGACACAGACATCCCGTCTCT	3180
Qy	3181	CCAGGGGCTGAGCCTCGAGGTAAAGAGGCGCACAGCCTGGCCCTCGTACGTAGCGAGTGG	3240
Db	3181	CCAGGGGCTGAGCCTCGAGGTAAAGAGGCGCACAGCCTGGCCCTCGTACGTAGCGAGTGG	3240
Qy	3241	CTGAGGGAAGAGCACAGTTGGTTCAGTCCCTGAGGCGCTTCATACCCCTTGGCTGTTTC	3300
Db	3241	CTGAGGGAAGAGCACAGTTGGTTCAGTCCCTGAGGCGCTTCATACCCCTTGGCTGTTTC	3300
Qy	3301	AGTGCTAATTGATGGCAAAAGATAAAGCACCTGATATCTCAGTGGCTCCGAGCACACCT	3360
Db	3301	AGTGCTAATTGATGGCAAAAGATAAAGCACCTGATATCTCAGTGGCTCCGAGCACACCT	3360
Qy	3361	GGGATGCTGTCTCAGAGAGCCATCCCTGTTTACCTGACAGATTTCCGAGAAACATTGCTTA	3420
Db	3361	GGGATGCTGTCTCAGAGAGCCATCCCTGTTTACCTGACAGATTTCCGAGAAACATTGCTTA	3420
Qy	3421	TGAGAGCAACAGCCGGGCTCGTATCACATGAGAGATTTCAGAGCAGCCAAAGAGGCCAA	3480
Db	3421	TGAGAGCAACAGCCGGGCTCGTATCACATGAGAGATTTCAGAGCAGCCAAAGAGGCCAA	3480
Qy	3481	CATACACACTTATTCGAGACACTCCCTGAGAAATACAAACCCAGAGTAGAGACAAAGG	3540
Db	3481	CATACACACTTATTCGAGACACTCCCTGAGAAATACAAACCCAGAGTAGAGACAAAGG	3540
Qy	3541	AACCCAGCTCTCTGTTGGGCCAGAAAGCGCATTTGCCATAGCTGGCTCTTTGTTAGACA	3600
Db	3541	AACCCAGCTCTCTGTTGGGCCAGAAAGCGCATTTGCCATAGCTGGCTCTTTGTTAGACA	3600
Qy	3601	GCCCTATATTTTGGCTTTTGGATGATAGCTATCATCAGTCTCGGATTAAGAAATGTAAAGGT	3660
Db	3601	GCCCTATATTTTGGCTTTTGGATGATGATCAGTCTCGGATTAAGAAATGTAAAGGT	3660
Qy	3661	TGTCCAAAGACCCCTGGAGCAAAAGCGCAGAGAGGCCCGCACCTCATTTGATCGCCACCG	3720
Db	3661	TGTCCAAAGACCCCTGGAGCAAAAGCGCAGAGAGGCCCGCACCTCATTTGATCGCCACCG	3720
Qy	3721	CTTTGTCCACATCCAGATGACATTTAATATGTGTGTTTCAGAATGGCCAAAGTCGAAGA	3780
Db	3721	CTTTGTCCACATCCAGATGACATTTAATATGTGTGTTTCAGAATGGCCAAAGTCGAAGA	3780
Qy	3781	GGATGGCACACATCAACAGCTGCTGGGCCCAAGAAAGCATCTATTTTTCATGGTCAGTG	3840
Db	3781	GGATGGCACACATCAACAGCTGCTGGGCCCAAGAAAGCATCTATTTTTCATGGTCAGTG	3840
Qy	3841	CCAGGCTGAGCAAAAGCCCTAGTACATGTGGCCATATGAGCTGTTAAATATTTTATAT	3900
Db	3841	CCAGGCTGAGCAAAAGCCCTAGTACATGTGGCCATATGAGCTGTTAAATATTTTATAT	3900
Qy	3901	ATTTCGTTAAACATGAGCATTTATCAAGATTAAAGGTGAGCACTTACTGTGAAAAACT	3960
Db	3901	ATTTCGTTAAACATGAGCATTTATCAAGATTAAAGGTGAGCACTTACTGTGAAAAACT	3960
Qy	3961	ATGTAGAACTACCTGTTTAACTTTCTTGCTGCAACTGAAGATCATTTCCACCAACTTAG	4020
Db	3961	ATGTAGAACTACCTGTTTAACTTTCTTGCTGCAACTGAAGATCATTTCCACCAACTTAG	4020
Qy	4021	AGCTTCACATTTTAAATTTAAAGAACCAAAACAAACATTTATCTGATGGAATTAATAATC	4080
Db	4021	AGCTTCACATTTTAAATTTAAAGAACCAAAACAAACATTTATCTGATGGAATTAATAATC	4080

QY	4081	TGGTGGTAATTCATTAATAAATTATAGACTAATTCAAAGTACGATTTGTTAATAATG	4140
Db	4081	TGGTGGTAATTCATTAATAAATTATAGACTAATTCAAAGTACGATTTGTTAATAATG	4140
QY	4141	TATATATTTGGTTATATATTTTATTTGTAACCTCGCTGCTGTAAGAAGATTATAGAAGTG	4200
Db	4141	TATATATTTGGTTATATATTTTATTTGTAACCTCGCTGCTGTAAGAAGATTATAGAAGTG	4200
QY	4201	GTAATAAGTACTGAATGTTTGGATTAAGTCGTACGCTATTAATAAAGTAACTTTTATAC	4260
Db	4201	GTAATAAGTACTGAATGTTTGAATTAAGTCGTACGCTATTAATAAAGTAACTTTTATAC	4260
QY	4261	AAAAAAAAAAAAAAAAAAAA 4279	
Db	4261	AAAAAAAAAAAAAAAAAAAA 4279	
RESULT 3			
AX105080		4279 bp	DNA
LOCUS			
DEFINITION	AX105080	4279 bp	DNA
ACCESSION	AX105080		
VERSION	AX105080.1	GI:13921230	
KEYWORDS			
SOURCE	dog.		
ORGANISM	Canis familiaris		
REFERENCE	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Carnivora; Fissipedia; Canidae; Canis.		
AUTHORS	1 (bases 1 to 4279) Stocker,P.J., Steimel-Crespi,D.T., Crespi,C.L., Relf,T.C. and Patten,C.J.		
TITLE	P-glycoproteins and uses thereof		
JOURNAL	Patent: WO 0123540-A 24 05-APR-2001;		
FEATURES	LOCATION/Qualifiers		
source	1..4279		
	/organism="Canis familiaris"		
	/db_xref="taxon:9615"		
	17..3862		
CDS	/note="unnamed protein product"		
	/codon_start=1		
	/protein_id="CAC37722.1"		
	/db_xref="gi:13921231"		
	/translation="MDPEGRKGSAAENFNMGRKSKKEKKKPTVSTFAMFRSN		
	WLDRLYMEVGTMAAIIHGAAALPLMMVLVGNMTDSFANAAGISRNKTFEVIINESITNNT		
	CHFTHLEBEMWTAYAYVYSGAGVIAVAAIQVSPFCIAAGROLILRKQFPAIMQ		
	ELGWPDVHDVGEVLTNRLFDVSKINEGCKMGKMGFPGISATFFGFLVGTGKMLTL		
	VTLASVPLGSLAATWAKILSFSTDKELATLAKGANADEVILAAIPVIAFGOKKGL		
	EYNNNLNEAKGIGIKKATITANISIGAFLITIASIALATPYGSLVLSSEYSICGL		
	TYEFSVLIGAFSIGQASPIEAFANARCAAAVEIKIIDNKPISDYSKSGHPRONIG		
	NIEFNHVSYPDSKREVKILKGLMLKVGSGCTVALVNGSGKSTVQLMORLDPD		
	GNVCJIDGDIRITINRHLREITGVVSOBPVLAFTIENIRYGRKENTMDEIEKAVE		
	ANATYPIKLPLPKRPETLVGSRGAOLSGGOKORIRIARGLVNPKLILLDLSALDE		
	SAVAVVOALDKARKGRTTIVIAHLRSTYRANMAYAGRDGVYVKGHDELMKKGTY		
	FVLVMOQRKNEIELENAETGKESSDYLENSPDDSSSLKRSPTRRSTIIPRITIGI		
	KIGTEDEENENENPVPSFWRILKLNSTENPVVQVIFCALINGGLQPARSTIIPRITIGI		
	FRDPEDETRKQNSNMFSVLFLVGLGISFIFFLQGFPGKAGELITRKRLVFNFSM		
	LIODVSMDDPKNTGATLTRLANDAAVKAISRLAIVITQNTALNGTGIILSLYIG		
	NOTLLITLAIPTIITAGVEMKMLSGOALKDKKELEBAGKATGEALENFTVSLTR		
	EOKFETRNQSIQVYKXNSIRKAHIFGYSFSTIVQNMVTFSGAFGAYLVANEFNNK		
	FQDVLVETSAIVFGAMAVGQVSPAPDYAKKAVSAHYIMIIKESPLIDSYPLGLPKR		
	NLEGVNTEVEVENVPTPRDIPVLQGLSLEKKQTLATVSGSGCKRSTVOLLERR		
	YPLKAGSVLIDKEIKHLNVMRLRAHLDIVSOEPLTFDCSIAENIATGDNRSYVSHEE		
	IOAAKEANIIHFIETLEPEKNTVRGDGTOLSGGOKORIRIARLVAPOHILIDERA		
	TSALDESEKVVQVQALDKRAREGRTCIIVIAHLRSTIQNADLLIVFQNKVKYEHGTHQD		
	LAQKGIYFSVSVQAGAKR"		
BASE COUNT	1296 a	833 c	1008 g 1142 t
ORIGIN			
Query Match	99.9%;	Score 4275.8;	DB 6; Length 4279;
Best Local Similarity	100.0%;	Pred. No. 0;	
Matches 4277; Conservative	0;	Mismatches	2; Indels 0; Gaps 0;

Qy	1	GGAGCGGAGGTCGGGATGGATCCGGAAGAGAGCGCGTTAAGGGGAGTGCAGAGAAACTT	60
Db	1	GGAGCGCGAGGTCGGGATGGATCCTGAAGAGAGCGCGTTAAGGGAGTGCAGAGAAACTT	60
Qy	61	CTGAAATGSGGCACAAAAAAGTAAAAAAATGAGAGAGAAAAAGAAACCAACTCTCAG	12
Db	61	CTGAAATGSGGCACAAAAAAGTAAAAAAAGAGAGAGAAAAAGAAACCAACTCTCAG	12
Qy	121	CACGTTTGCATGTTTCGCTATCCATAAATTTGGCTTGATAGGTTGTATATGTTGGTGGAC	18
Db	121	CACGTTTGCATGTTTCGCTATCCATAAATTTGGCTTGATAGGTTGTATATGTTGGTGGAC	18
Qy	181	AATGGGTGCATCAATCCATGAGAGCTGCACCTCCTCATGATGCTGGTTTGGAAACAT	24
Db	181	AATGGGTGCATCAATCCATGAGAGCTGCACCTCCTCATGATGCTGGTTTGGAAACAT	24
Qy	241	GACAGATAGCTTTCGAATGAGAGATTTCAAAAAACAACCTTTCCAGTTATATTTAA	30
Db	241	GACAGATAGCTTTCGAATGAGAGATTTCAAAAAACAACCTTTCCAGTTATATTTAA	30
Qy	301	TGAAGTATTTACGAACATTTCAACAATTTCAACCATCTGAGAGAGAAATGACAC	36
Db	301	TGAAGTATTTACGAACATTTCAACAATTTCAACCATCTGAGAGAGAAATGACAC	36
Qy	361	GTATGCTATTTATACAGTGGATGCGTGGCGGCGTGGTGGTCTTCATCCAGT	42
Db	361	GTATGCTATTTATACAGTGGATGCGTGGCGGCGTGGTGGTCTTCATCCAGT	42
Qy	421	TTCAATTCGGGCGCTGGCAGCAGGAAGACAGATACTCAAAATTTAGAAAACAATTTTCA	48
Db	421	TTCAATTCGGGCGCTGGCAGCAGGAAGACAGATACTCAAAATTTAGAAAACAATTTTCA	48
Qy	481	TGCTATCATCGCAGCAGAGATTTGGCGTTTACGTGATAGTACGTTGGGAGCTTAAAC	54
Db	481	TGCTATCATCGCAGCAGAGATTTGGCGTTTACGTGATAGTACGTTGGGAGCTTAAAC	54
Qy	541	CCGGCTACGAGCAGATGTCCTCAAAATCAATGAAGAAATTTGGCACAATTTGGAATGTT	60
Db	541	CCGGCTACGAGCAGATGTCCTCAAAATCAATGAAGAAATTTGGCACAATTTGGAATGTT	60
Qy	601	CTTTCACTCAATAGCAACAATTTTTCACGGTTTATATGATGGGGTTTACAGTGGTGA	66
Db	601	CTTTCACTCAATAGCAACAATTTTTCACGGTTTATATGATGGGGTTTACAGTGGTGA	66
Qy	661	GCATACCTTGATTTTGGCCATCAGCCCTGTTCTTGAGCTTTCAGCCGCATCTGGGC	72
Db	661	GCATACCTTGATTTTGGCCATCAGCCCTGTTCTTGAGCTTTCAGCCGCATCTGGGC	72
Qy	721	AAAGATACTATTTTATTACTGATTAAGAAACCTTGGCCTATGCAAAAGCTGAGCAGT	78
Db	721	AAAGATACTATTTTATTACTGATTAAGAAACCTTGGCCTATGCAAAAGCTGAGCAGT	78
Qy	781	AACTGAAGAAGTCTTAGCAGCAATCAGAACTGATTTGGCTTTGGAGGACAAAAGAAAG	84
Db	781	AACTGAAGAAGTCTTAGCAGCAATCAGAACTGATTTGGCTTTGGAGGACAAAAGAAAG	84
Qy	841	ACTTGAAGGATCAACAACAATTTTGAAGAGAGTAAAGAAATTTGGATTAAGAAAGCTAT	90
Db	841	ACTTGAAGGATCAACAACAATTTTGAAGAGAGTAAAGAAATTTGGATTAAGAAAGCTAT	90
Qy	901	CACGGCCAACTTTCTATGTTGGTCCGCTTCTTATGATCTATGATATATATGCTGGC	96
Db	901	CACGGCCAACTTTCTATGTTGGTCCGCTTCTTATGATCTATGATATATATGCTGGC	96
Qy	961	TTTCTGGATGGAGACCTTCCTGCTCCAGTGAATTTCTATGACAAAGTACTCAC	102
Db	961	TTTCTGGATGGAGACCTTCCTGCTCCAGTGAATTTCTATGACAAAGTACTCAC	102
Qy	1021	TGCTCTTTTCTGATTTATTTAGGGCTTTTGTATTTGAGAGGCAATCCCAAGCATTTGA	108
Db	1021	TGCTCTTTTCTGATTTATTTAGGGCTTTTGTATTTGAGAGGCAATCCCAAGCATTTGA	108

QY	1081	AGCATTTGCAAAAGCAAGAGAGACAGCTTATGAAATCTTTCAGATATATGACATTAAC	1140
Db	1081	AGCATTTGCAAAAGCAAGAGAGAGACAGCTTATGAAATCTTTCAGATATATGACATTAAC	1140
QY	1141	AAGATTCACAGTATATTCGAAGTGGACATTAACCAAGATATATTAAGGAAATTGGA	1200
Db	1141	AAGCATTTGACAGTATATTCGAAGAGTGGACATTAACCAAGATATATTAAGGAAATTGGA	1200
QY	1201	ATTCAAAAATTTGTTCATCTCAGTTTACCTCTTCGAAAAGAGTTAAATCTTAAAGGCTC	1260
Db	1201	ATTCAAAAATTTGTTCATCTCAGTTTACCTCTTCGAAAAGAGTTAAATCTTAAAGGCTC	1260
QY	1261	CAACCTGAAGGTTTCAGAGTGGGCGACAGTGGCGCTGTTGGGACAGTGGCGGAA	1320
Db	1261	CAACCTGAAGGTTTCAGAGTGGGCGACAGTGGCGCTGTTGGGACAGTGGCGGAA	1320
QY	1321	GAGCAGCAGCCGTGCTGTTGTTGGCCACCAAGTAGTGSAAACATTCGCTATGCCCCGA	1380
Db	1321	GAGCAGCAGCCGTGCTGTTGTTGGCCACCAAGTAGTGSAAACATTCGCTATGCCCCGA	1380
QY	1381	TGATGGACAGACATTTAGGACATTAATTTGAAGGCATCTTCGGGAAATTTACGTGTGAT	1440
Db	1381	TATGAGACAGACATTTAGGACATTAATTTGAAGGCATCTTCGGGAAATTTACGTGTGAT	1440
QY	1441	GAGTCAGAGCCTGTGTTGTTGGCCACCAAGTAGTGSAAACATTCGCTATGCCCCGA	1500
Db	1441	GAGTCAGAGCCTGTGTTGTTGGCCACCAAGTAGTGSAAACATTCGCTATGCCCCGA	1500
QY	1501	AAAGTCCACCATGGATCAGATTAGAAAGCTGTTTAAGGAAGCCAAATGCCATTTAT	1560
Db	1501	AAAGTCCACCATGGATCAGATTAGAAAGCTGTTTAAGGAAGCCAAATGCCATTTAT	1560
QY	1561	CATGAACCTACTATTAATTTGACCTCTGTGTTGGAGAGAGAGGCCCCAGCTGAGTGG	1620
Db	1561	CATGAACCTACTATTAATTTGACCTCTGTGTTGGAGAGAGAGGCCCCAGCTGAGTGG	1620
QY	1621	TGSCACAAAACAGAAATGCCATTGCTGGGCCCTGGTTTGGCAACCCCAAGATTCCTCT	1680
Db	1621	TGSCACAAAACAGAAATGCCATTGCTGGGCCCTGGTTTGGCAACCCCAAGATTCCTCT	1680
QY	1681	GCTGGATGAGGCAACGCTCAGCTCTGTGGACACGTGAAGACAGTGGTTAGGTGGCCCT	1740
Db	1681	GCTGGATGAGGCAACGCTCAGCTCTGTGGACACGTGAAGACAGTGGTTAGGTGGCCCT	1740
QY	1741	GGATTAAGGCCAAGAAAGGCCGGACTACCATTTGTATAGCTCATTCGTTTGTCTACGATTG	1800
Db	1741	GGATTAAGGCCAAGAAAGGCCGGACTACCATTTGTATAGCTCATTCGTTTGTCTACGATTG	1800
QY	1801	TAATGGCGATGTCATCTGCGTGTGTTGATGATGAGGTATGTTGTGGAGAAAGAAATCATGA	1860
Db	1801	TATGCGCATGTCATCTGCGTGTGTTGATGATGAGGTATGTTGTGGAGAAAGAAATCATGA	1860
QY	1861	TGACATCATGAAGAAGAGGCGATTTACTTCAAACTTGTCACAAATGCACAGCAAGAGSAA	1920
Db	1861	TGACATCATGAAGAAGAGGCGATTTACTTCAAACTTGTCACAAATGCACAGCAAGAGSAA	1920
QY	1921	TGAAAATTTGATAGAAAATGCGCATGCGTGTGATCCAAAAGTGAAGTGAATGCCTTGGAAT	1980
Db	1921	TGAAAATTTGATAGAAAATGCGCATGCGTGTGATCCAAAAGTGAAGTGAATGCCTTGGAAT	1980
QY	1981	GTTCTCAAAAAGATTCAAGGTTCCAGTTTAAATTAATAAAGATACCTCCAGAGAGATATCA	2040
Db	1981	GTTCTCAAAAAGATTCAAGGTTCCAGTTTAAATTAATAAAGATACCTCCAGAGAGATATCA	2040
QY	2041	TGCACCAACAAGGCCAACAAGAAAGCTTGGTACAAAAGAGACTTGAATGAGATGTAC	2100
Db	2041	TGCACCAACAAGGCCAACAAGAAAGCTTGGTACAAAAGAGACTTGAATGAGATGTAC	2100
QY	2101	TCCAGTTTCTTTCGAGGATTTCTGAAGCTGAACATCACTAATAGGCTTATTTTGTGGT	2160
Db	2101	TCCAGTTTCTTTCGAGGATTTCTGAAGCTGAACATCACTAATAGGCTTATTTTGTGGT	2160
QY	2161	TGGTATATTTTGTGCTATTTATAACGAGGCGCTGCACACGACATTTTCAATTAATATTTTC	2220

|||||
Db 2161 TGGTATATTTGGCTATTAATTAAGGAGGCGCTCAACACATATTTCAATATATTTTC 2220
OY 2221 AAGGATTAATAGGATCTTACCCGAGATGAGATCTCGAACAACAAAGACAGATATAA 2280
Db 2221 AAGGATTAATAGGATCTTACCCGAGATGAGATCTCGAACAACAAAGACAGATATAA 2280
OY 2281 CATGTTTCTGATGTTTCTAGTCTTGAATATTTCTTTATTTACATTTTCTTCCA 2340
Db 2281 CATGTTTCTGATGTTTCTAGTCTTGAATATTTCTTTATTTACATTTTCTTCCA 2340
OY 2341 GGGCTTCACTTTTGGCAAGCTGGGAGATCTCTCAAGGGCGTTCATCATGTTTT 2400
Db 2341 GGGCTTCACTTTTGGCAAGCTGGGAGATCTCTCAAGGGCGTTCATCATGTTTT 2400
OY 2401 CAGATCCATGCTGAGACAGATGTCAGTGTGATGATGACCTTAAACACACATGGAGC 2460
Db 2401 CAGATCCATGCTGAGACAGATGTCAGTGTGATGATGACCTTAAACACACATGGAGC 2460
OY 2461 ATTGACAAACGAGCTTGCCAAATGATCGGCTCAAGTTAAAGGGCTATAGTTCCAGCT 2520
Db 2461 ATTGACAAACGAGCTTGCCAAATGATCGGCTCAAGTTAAAGGGCTATAGTTCCAGCT 2520
OY 2521 TGTCTGATTAACCCAGATATAGCAATCTTGGACAGGCATTTATATTCCTTAATCTA 2580
Db 2521 TGTCTGATTAACCCAGATATAGCAATCTTGGACAGGCATTTATATTCCTTAATCTA 2580
OY 2581 TGGTGGCAATTAACACTTTTACTCTTAAATTTGTAATCCATTCATTCAGAGAGT 2640
Db 2581 TGGTGGCAATTAACACTTTTACTCTTAAATTTGTAATCCATTCATTCAGAGAGT 2640
OY 2641 TGTGTAATGAAATGTTGTCAGCAAGCAGTGAAGATGAAGAGAGCTAGAGAGC 2700
Db 2641 TGTGTAATGAAATGTTGTCAGCAAGCAGTGAAGATGAAGAGAGCTAGAGAGC 2700
OY 2701 TGGGAAGATGCTACAGAGCCATCGAAATCTCCGAACTGTCTTCTTGGCTGGGA 2760
Db 2701 TGGGAAGATGCTACAGAGCCATCGAAATCTCCGAACTGTCTTCTTGGCTGGGA 2760
OY 2761 GCAGAAATTTGAATCATGTATGACAGAGTTGCAAGTACCATTCAGAAACTCTTGGAG 2820
Db 2761 GCAGAAATTTGAATCATGTATGACAGAGTTGCAAGTACCATTCAGAAACTCTTGGAG 2820
OY 2821 GAAAGACACATCTTGGGGTCTCATTTCTATCACCCAGCATGATGATTTTCTCTA 2880
Db 2821 GAAAGACACATCTTGGGGTCTCATTTCTATCACCCAGCATGATGATTTTCTCTA 2880
OY 2881 TGTGCTGCTTTCGGGTTGGTGTCTACTGTGTGCAATGATGATGAACTTTCAGGA 2940
Db 2881 TGTGCTGCTTTCGGGTTGGTGTCTACTGTGTGCAATGATGATGAACTTTCAGGA 2940
OY 2941 TGTCTTCTTGTATTCACGATTTGTCTTGGTGCATGGAGAGGGGCAAGTCTTC 3000
Db 2941 TGTCTTCTTGTATTCACGATTTGTCTTGGTGCATGGAGAGGGGCAAGTCTTC 3000
OY 3001 ATTTGCTCCGATATGCCAAAGCCAAAGTATCAGAGCCCACTCATCATGATCATGGA 3060
Db 3001 ATTTGCTCCGATATGCCAAAGCCAAAGTATCAGAGCCCACTCATCATGATCATGGA 3060
OY 3061 AAAAAGCCTCTGATTCACGACTACAGCCCTCAGGCTCAAGCCAAATACGTTGGAAGG 3120
Db 3061 AAAAAGCCTCTGATTCACGACTACAGCCCTCAGGCTCAAGCCAAATACGTTGGAAGG 3120
OY 3121 AAATGTGACATTTATAGAGTCTGTTGCAACTATCCACTCGACAGACATCCCGTGT 3180
Db 3121 AAATGTGACATTTATAGAGTCTGTTGCAACTATCCACTCGACAGACATCCCGTGT 3180
OY 3181 CCAGGGGCTGAGCTCGAGGTGAAGAGGCGCAGAGCGTGGCCCTGTAGTGAAGCTGG 3240
Db 3181 CCAGGGGCTGAGCTCGAGGTGAAGAGGCGCAGAGCGTGGCCCTGTAGTGAAGCTGG 3240
OY 3241 CTGTGGAGAGCACAGTTGTCAGTCTCTAGAGCGCTTATGACCCCTTGGCTGTTTC 3300
|||||

Db 3241 CTGTGGAGAGCACAGTTGTTCACTCTAGAGCGCTTCTATGACCCCTTGGCTGTTTC 3300
OY 3301 AGTGTATTTGATGGCAAGATTAAGCACTGATGTCAGTGGCTCCGAGCACACCT 3360
Db 3301 AGTGTATTTGATGGCAAGATTAAGCACTGATGTCAGTGGCTCCGAGCACACCT 3360
OY 3361 GGGCATGCTGTCAGAGCCCTCTCTTGTGATGAGCATGAGATTCGGAACAATTCGCTA 3420
Db 3361 GGGCATGCTGTCAGAGCCCTCTCTTGTGATGAGCATGAGATTCGGAACAATTCGCTA 3420
OY 3421 TGGAGCAACAGCCGAGTGTATCATGATGAAGATTAATGACGACCCCAAGAGGCCAA 3480
Db 3421 TGGAGCAACAGCCGAGTGTATCATGATGAAGATTAATGACGACCCCAAGAGGCCAA 3480
OY 3481 CATACACACTTCATCGAGACACTCCCTGAGAAATATCAACACAGAGTAGAGCAAAAG 3540
Db 3481 CATACACACTTCATCGAGACACTCCCTGAGAAATATCAACACAGAGTAGAGCAAAAG 3540
OY 3541 AACCCAGCTCTGAGTGGGCAAGAACAGGCAATTCGATGCTGCTCTTGTATGACA 3600
Db 3541 AACCCAGCTCTGAGTGGGCAAGAACAGGCAATTCGATGCTGCTCTTGTATGACA 3600
OY 3601 GCTCATATTTTCTTTTGGATGAGCTACATCACTCTGATGATCAGAAAGTGAAGAGT 3660
Db 3601 GCTCATATTTTCTTTTGGATGAGCTACATCACTCTGATGATCAGAAAGTGAAGAGT 3660
OY 3661 TGTCCAGAAAGCCCTGGACAAACAGAGAGGCCCATCTCATGTTGATGCTCCACCG 3720
Db 3661 TGTCCAGAAAGCCCTGGACAAACAGAGAGGCCCATCTCATGTTGATGCTCCACCG 3720
OY 3721 CTGTCCACCATCCAGATGACATTTATAGTGTGTTTCAATGGAATGGAAGTCAAGGA 3780
Db 3721 CTGTCCACCATCCAGATGACATTTATAGTGTGTTTCAATGGAATGGAAGTCAAGGA 3780
OY 3781 GCATGAGCAACATCAACAGCTGCTGCCCAAGAAAGCATATTTTCCATGCTCAGT 3840
Db 3781 GCATGAGCAACATCAACAGCTGCTGCCCAAGAAAGCATATTTTCCATGCTCAGT 3840
OY 3841 CCAGGCTGAGCAAGCGCTATGTAACATGTCGCTGATGAGCTGTTAAATTTTAAAT 3900
Db 3841 CCAGGCTGAGCAAGCGCTATGTAACATGTCGCTGATGAGCTGTTAAATTTTAAAT 3900
OY 3901 ATTTGTTTAAACATGSCATTTATCAAGTTTAAAGGTGAGCACTTACTGAAAAACT 3960
Db 3901 ATTTGTTTAAACATGSCATTTATCAAGTTTAAAGGTGAGCACTTACTGAAAAACT 3960
OY 3961 ATGTAGAACTACTGTTTAACTTTCTTGTGCAAGTGAAGATCATCCCAAGTTCAG 4020
Db 3961 ATGTAGAACTACTGTTTAACTTTCTTGTGCAAGTGAAGATCATCCCAAGTTCAG 4020
OY 4021 AGTCTCAGATTTTAAATTAAGAACCAAGAAACATTTCTGATGATGAAATATAC 4080
Db 4021 AGTCTCAGATTTTAAATTAAGAACCAAGAAACATTTCTGATGATGAAATATAC 4080
OY 4081 TGTGTTAAATGCAATTAATAATTTATAGAGTAATCAAGATGATTTGTTAATAATG 4140
Db 4081 TGTGTTAAATGCAATTAATAATTTATAGAGTAATCAAGATGATTTGTTAATAATG 4140
OY 4141 TATAATTTTGTATATTTTATTTGTAACCTACTGCTTGTGTAAGATTAATGAAAGTG 4200
Db 4141 TATAATTTTGTATATTTTATTTTGTACTACTGCTTGTGTAAGATTAATGAAAGTG 4200
OY 4201 GTAAAAAGTACTGATGTTGAATTAAGGCTAGCTATTAATTAACCTTATATTC 4260
Db 4201 GTAAAAAGTACTGATGTTGAATTAAGGCTAGCTATTAATTAACCTTATATTC 4260
OY 4261 AAAAAAAAAAAAAAAAAA 4279
Db 4261 AAAAAAAAAAAAAAAAAA 4279

RESULT 4
AA105082

CDS

BASE COUNT	1296	a	833	c	1009	g	1141	t
ORIGIN								

Query Match

Best Local Similarity	score	DB 6;	Length
99.96;	4272.6;		4279;
99.98;			
Pred. No. 0;			

	Matches	4275;	Conservative	0;	Mismatches	0;
	Pred.	No.	0;			

0; Mismatches 4; Indels 0; Gaps 0;

Y 1 GGAGCGGAGGTCGGGATGCATCCTGAAGGAGGCCCTAAGGGGAGTGCAGAGACGACCTTTC 60

1 GCAACCCCAAGTGCAGAGAGAACTT 60

1 GGAGCCGAGGTCGGATGATCCTGAAGGAGGCCCGTAAGGGAGTGCAGAGAACTT 60

61 CTGGAAATGGCCCAAAATCTT..... 60

01 CTGGAAATGGGCAAAAAGTAAAAAAATGAGAAGAAAGAAAGAAACCACTGTCA 120

61 CTGGAATGGCCAAAACCTATAAAGCATTTCCTTGTGC
CG

120

01 C TGGGAAATGGGCAAAAAGTAAAAAAGAGAGAGAAAGAAAGAAAGAAACCAACTGTCTAG 120

121 CACGTTGCAATGTTCCGTATTCAAATTGCCCTTCATAAGCCTGGTAA
120

|||||GCTATTCGAATGGCTTGATATATGTTGGTGGGAC 180

121 CACGTTGCATGTTCGCTATTCAATTGGCTTAGCTCTATACTCCCTCCT

101
102
103
104
105
106
107
108
109
110
111
112
113
114
115
116
117
118
119
120
121
122
123
124
125
126
127
128
129
130
131
132
133
134
135
136
137
138
139
140
141
142
143
144
145
146
147
148
149
150
151
152
153
154
155
156
157
158
159
160
161
162
163
164
165
166
167
168
169
170
171
172
173
174
175
176
177
178
179
180
181
182
183
184
185
186
187
188
189
190
191
192
193
194
195
196
197
198
199
200
201
202
203
204
205
206
207
208
209
210
211
212
213
214
215
216
217
218
219
220
221
222
223
224
225
226
227
228
229
230
231
232
233
234
235
236
237
238
239
240
241
242
243
244
245
246
247
248
249
250
251
252
253
254
255
256
257
258
259
260
261
262
263
264
265
266
267
268
269
270
271
272
273
274
275
276
277
278
279
280
281
282
283
284
285
286
287
288
289
290
291
292
293
294
295
296
297
298
299
300
301
302
303
304
305
306
307
308
309
310
311
312
313
314
315
316
317
318
319
320
321
322
323
324
325
326
327
328
329
330
331
332
333
334
335
336
337
338
339
340
341
342
343
344
345
346
347
348
349
350
351
352
353
354
355
356
357
358
359
360
361
362
363
364
365
366
367
368
369
370
371
372
373
374
375
376
377
378
379
380
381
382
383
384
385
386
387
388
389
390
391
392
393
394
395
396
397
398
399
400
401
402
403
404
405
406
407
408
409
410
411
412
413
414
415
416
417
418
419
420
421
422
423
424
425
426
427
428
429
430
431
432
433
434
435
436
437
438
439
440
441
442
443
444
445
446
447
448
449
450
451
452
453
454
455
456
457
458
459
460
461
462
463
464
465
466
467
468
469
470
471
472
473
474
475
476
477
478
479
480
481
482
483
484
485
486
487
488
489
490
491
492
493
494
495
496
497
498
499
500
501
502
503
504
505
506
507
508
509
510
511
512
513
514
515
516
517
518
519
520
521
522
523
524
525
526
527
528
529
530
531
532
533
534
535
536
537
538
539
540
541
542
543
544
545
546
547
548
549
550
551
552
553
554
555
556
557
558
559
560
561
562
563
564
565
566
567
568
569
570
571
572
573
574
575
576
577
578
579
580
581
582
583
584
585
586
587
588
589
590
591
592
593
594
595
596
597
598
599
600
601
602
603
604
605
606
607
608
609
610
611
612
613
614
615
616
617
618
619
620
621
622
623
624
625
626
627
628
629
630
631
632
633
634
635
636
637
638
639
640
641
642
643
644
645
646
647
648
649
650
651
652
653
654
655
656
657
658
659
660
661
662
663
664
665
666
667
668
669
670
671
672
673
674
675
676
677
678
679
680
681
682
683
684
685
686
687
688
689
690
691
692
693
694
695
696
697
698
699
700
701
702
703
704
705
706
707
708
709
710
711
712
713
714
715
716
717
718
719
720
721
722
723
724
725
726
727
728
729
730
731
732
733
734
735
736
737
738
739
740
741
742
743
744
745
746
747
748
749
750
751
752
753
754
755
756
757
758
759
760
761
762
763
764
765
766
767
768
769
770
771
772
773
774
775
776
777
778
779
780
781
782
783
784
785
786
787
788
789
790
791
792
793
794
795
796
797
798
799
800
801
802
803
804
805
806
807
808
809
810
811
812
813
814
815
816
817
818
819
820
821
822
823
824
825
826
827
828
829
830
831
832
833
834
835
836
837
838
839
840
841
842
843
844
845
846
847
848
849
850
851
852
853
854
855
856
857
858
859
860
861
862
863
864
865
866
867
868
869
870
871
872
873
874
875
876
877
878
879
880
881
882
883
884
885
886
887
888
889
890
891
892
893
894
895
896
897
898
899
900
901
902
903
904
905
906
907
908
909
910
911
912
913
914
915
916
917
918
919

181 AATGGCTGCCATCATCCATGGAGCTGCACCTCCCTCTCATGATGCTGTGTTTTCGAACTT 340

181
ATTCCTGCTGCTTTTGGAAACAT 240

181 AATGGCTGCCATCATCGAGCTGCACTCCCTCATGATGCTGTTTGGAAACAT 240

241 GACAGATACCCTTCCGAAATTTCAGTTGGGAACAT 240

241 GACAGATAGCTTGGCAATGCAGGATTTCAAGAAACAACCTTTCCAGTATAATTAA 300

241 GACAGATGCTTGGCAATCCCGCCTTCCTTTTGCTGCTGCCTGTGGGTA
|||||
|||

241 GACAGATAGCTTTGCAAATGCAGGAATTTCAAGAACAAACTTTCCAGTTATAATTAA 300

000
001
002
003
004
005
006
007
008
009
010
011
012
013
014
015
016
017
018
019
020
021
022
023
024
025
026
027
028
029
030
031
032
033
034
035
036
037
038
039
040
041
042
043
044
045
046
047
048
049
050
051
052
053
054
055
056
057
058
059
060
061
062
063
064
065
066
067
068
069
070
071
072
073
074
075
076
077
078
079
080
081
082
083
084
085
086
087
088
089
090
091
092
093
094
095
096
097
098
099
100
101
102
103
104
105
106
107
108
109
110
111
112
113
114
115
116
117
118
119
120
121
122
123
124
125
126
127
128
129
130
131
132
133
134
135
136
137
138
139
140
141
142
143
144
145
146
147
148
149
150
151
152
153
154
155
156
157
158
159
160
161
162
163
164
165
166
167
168
169
170
171
172
173
174
175
176
177
178
179
180
181
182
183
184
185
186
187
188
189
190
191
192
193
194
195
196
197
198
199
200
201
202
203
204
205
206
207
208
209
210
211
212
213
214
215
216
217
218
219
220
221
222
223
224
225
226
227
228
229
230
231
232
233
234
235
236
237
238
239
240
241
242
243
244
245
246
247
248
249
250
251
252
253
254
255
256
257
258
259
260
261
262
263
264
265
266
267
268
269
270
271
272
273
274
275
276
277
278
279
280
281
282
283
284
285
286
287
288
289
290
291
292
293
294
295
296
297
298
299
300
301
302
303
304
305
306
307
308
309
310
311
312
313
314
315
316
317
318
319
320
321
322
323
324
325
326
327
328
329
330
331
332
333
334
335
336
337
338
339
340
341
342
343
344
345
346
347
348
349
350
351
352
353
354
355
356
357
358
359
360
361
362
363
364
365
366
367
368
369
370
371
372
373
374
375
376
377
378
379
380
381
382
383
384
385
386
387
388
389
390
391
392
393
394
395
396
397
398
399
400
401
402
403
404
405
406
407
408
409
410
411
412
413
414
415
416
417
418
419
420
421
422
423
424
425
426
427
428
429
430
431
432
433
434
435
436
437
438
439
440
441
442
443
444
445
446
447
448
449
450
451
452
453
454
455
456
457
458
459
460
461
462
463
464
465
466
467
468
469
470
471
472
473
474
475
476
477
478
479
480
481
482
483
484
485
486
487
488
489
490
491
492
493
494
495
496
497
498
499
500
501
502
503
504
505
506
507
508
509
510
511
512
513
514
515
516
517
518
519
520
521
522
523
524
525
526
527
528
529
530
531
532
533
534
535
536
537
538
539
540
541
542
543
544
545
546
547
548
549
550
551
552
553
554
555
556
557
558
559
560
561
562
563
564
565
566
567
568
569
570
571
572
573
574
575
576
577
578
579
580
581
582
583
584
585
586
587
588
589
590
591
592
593
594
595
596
597
598
599
600
601
602
603
604
605
606
607
608
609
610
611
612
613
614
615
616
617
618
619
620
621
622
623
624
625
626
627
628
629
630
631
632
633
634
635
636
637
638
639
640
641
642
643
644
645
646
647
648
649
650
651
652
653
654
655
656
657
658
659
660
661
662
663
664
665
666
667
668
669
670
671
672
673
674
675
676
677
678
679
680
681
682
683
684
685
686
687
688
689
690
691
692
693
694
695
696
697
698
699
700
701
702
703
704
705
706
707
708
709
710
711
712
713
714
715
716
717
718
719
720
721
722
723
724
725
726
727
728
729
730
731
732
733
734
735
736
737
738
739
740
741
742
743
744
745
746
747
748
749
750
751
752
753
754
755
756
757
758
759
760
761
762
763
764
765
766
767
768
769
770
771
772
773
774
775
776
777
778
779
780
781
782
783
784
785
786
787
788
789
790
791
792
793
794
795
796
797
798
799
800
801
802
803
804
805
806
807
808
809
810
811
812
813
814
815
816
817
818
8

OY 1381 TGATGACAGACATTTAGACCAATATGTATAGCACTCTTCGGGAATTTACTGTGTGCT 1440
|||||
Db 1381 TGATGACAGACATTTAGACCAATATGTATAGCACTCTTCGGGAATTTACTGTGTGCT 1440
OY 1441 GAGTCAGAGAGCCTGTGTGTTTGGCCACACGATAGCTGAAACATTTGGTATAGGCGCGCA 1500
|||||
Db 1441 GAGTCAGAGAGCCTGTGTGTTTGGCCACACGATAGCTGAAACATTTGGTATAGGCGCGCA 1500
OY 1501 AAATGTCCACATGATGATGTGAGAAAGCTGTTAAGGAAGCCAAATGCCATATGTTTAT 1560
|||||
Db 1501 AAATGTCCACATGATGATGTGAGAAAGCTGTTAAGGAAGCCAAATGCCATATGTTTAT 1560
OY 1561 CATGAACACTACCAATTAATTTGACACTCTGTGTTGAGAGAGAGGGCCCAAGCTGATG 1620
|||||
Db 1561 CATGAACACTACCAATTAATTTGACACTCTGTGTTGAGAGAGAGGGCCCAAGCTGATG 1620
OY 1621 TGACAGAAACAGACATGCGCATTTGCTCGGGCCCTGGTTGSCAACCCCAAGATTTCT 1680
|||||
Db 1621 TGACAGAAACAGACATGCGCATTTGCTCGGGCCCTGGTTGSCAACCCCAAGATTTCT 1680
OY 1681 GCTGGATGAGGCAAGCTCAGCTCTGACACATGAAAGTAAAGCATGTGTCAGGTGGCCCT 1740
|||||
Db 1681 GCTGGATGAGGCAAGCTCAGCTCTGACACATGAAAGTAAAGCATGTGTCAGGTGGCCCT 1740
OY 1741 GGATTAAGCCAGAAAGGCCGAGCTACCATGTGATAGCTCATGCTTTGTCTACAGTTTC 1800
|||||
Db 1741 GGATTAAGCCAGAAAGGCCGAGCTACCATGTGATAGCTCATGCTTTGTCTACAGTTTC 1800
OY 1801 TAATGCCGATGTCATTTGCTGTTTGTATGATGAGTGCATTTGTGAGAAAGAAATCATGA 1860
|||||
Db 1801 TAATGCCGATGTCATTTGCTGTTTGTATGATGAGTGCATTTGTGAGAAAGAAATCATGA 1860
OY 1861 TGAACCTCATGAAGAAGGAGCAATTTACTCAAACTTGTCAACATGACAGCAAGAGAAA 1920
|||||
Db 1861 TGAACCTCATGAAGAAGGAGCAATTTACTCAAACTTGTCAACATGACAGCAAGAGAAA 1920
OY 1921 TGAATTTAGTTAGAAATGCGCACTGTGTGAATCCAAAGTGAAGATGCCCTGGAAAT 1980
|||||
Db 1921 TGAATTTAGTTAGAAATGCGCACTGTGTGAATCCAAAGTGAAGATGCCCTGGAAAT 1980
OY 1981 GTCTCCAAAGATTTAGGCTCCAGTTTAATTAATTAAGAAAGATCACTCGAGAGATATGA 2040
|||||
Db 1981 GTCTCCAAAGATTTAGGCTCCAGTTTAATTAATTAAGAAAGATCACTCGAGAGATATGA 2040
OY 1981 GTCTCCAAAGATTTAGGCTCCAGTTTAATTAATTAAGAAAGATCACTCGAGAGATATGA 2040
|||||
Db 1981 GTCTCCAAAGATTTAGGCTCCAGTTTAATTAATTAAGAAAGATCACTCGAGAGATATGA 2040
OY 2041 TGACACCAAGGCCAAGACAGAAAGCTTGTATCAAAAGAGACTTGAATGAGAAATGACC 2100
|||||
Db 2041 TGACACCAAGGCCAAGACAGAAAGCTTGTATCAAAAGAGACTTGAATGAGAAATGACC 2100
OY 2101 TCCAGTTTCCTTCTGGAGATTTCTGAACCTGAACCTGAATGGCCTTATTTTGTGT 2160
|||||
Db 2101 TCCAGTTTCCTTCTGGAGATTTCTGAACCTGAACCTGAATGGCCTTATTTTGTGT 2160
OY 2161 TGGATATTTTGTCTATTAATAAGGAGGCTGCAACCAAGCATTTTCAATATTTTC 2220
|||||
Db 2161 TGGATATTTTGTCTATTAATAAGGAGGCTGCAACCAAGCATTTTCAATATTTTC 2220
OY 2221 AAGGATTAATAGGATCTTTAACCAGATAGAGATCTGAAACAAAGACAGAAATAGTAA 2280
|||||
Db 2221 AAGGATTAATAGGATCTTTAACCAGATAGAGATCTGAAACAAAGACAGAAATAGTAA 2280
OY 2281 CATGTTTTCTGTATTTTCTAGTCCCTGGAATTAATTTCTTTTATTAATTTTCTCCA 2340
|||||
Db 2281 CATGTTTTCTGTATTTTCTAGTCCCTGGAATTAATTTCTTTTATTAATTTTCTCCA 2340
OY 2341 GGGCTTTCACATTTGGCAAAGCTGGGAGATCTCCTACTAAGCGGCTTCATCATGTTT 2400
|||||
Db 2341 GGGCTTTCACATTTGGCAAAGCTGGGAGATCTCCTACTAAGCGGCTTCATCATGTTT 2400
OY 2401 CAGATCCATGCTGAGACAGGATGTCAGCTGTTTGAATGACCTTAAACACCACTGGAGC 2460
|||||
Db 2401 CAGATCCATGCTGAGACAGGATGTCAGCTGTTTGAATGACCTTAAACACCACTGGAGC 2460
OY 2461 ATTGACAAACAGGCTTGGCAATGATGGGCTCAAGTTAAAGGGCTATAGGTTCCAGGCT 2520
|||||

|||||
Db 2461 ATTGACAAACAGGCTTGGCAATGATGGGCTCAAGTTAAAGGGGCTATAGGTTCCAGGCT 2520
OY 2521 TGCTGTCATTACCCGAATATAGCAAAATCTTGGGACAGCAATTTATTTATCCTTATCTA 2580
|||||
Db 2521 TGCTGTCATTACCCGAATATAGCAAAATCTTGGGACAGCAATTTATTTATCTTATCTA 2580
OY 2581 TGGTGGCAATTAACACTTTTACTCTTAGCAATGTATGCCCTATTCGCAATPACAGAGAT 2640
|||||
Db 2581 TGGTGGCAATTAACACTTTTACTCTTAGCAATGTATGCCCTATTCGCAATPACAGAGAT 2640
OY 2641 TGTGGAATGAATATGTGTCTGACAAAGCACTGAAAGATTAAGAAAGCTAGAGAGAGC 2700
|||||
Db 2641 TGTGGAATGAATATGTGTCTGACAAAGCACTGAAAGATTAAGAAAGCTAGAGAGAGC 2700
OY 2701 TGGGAAATGTCTACAGAAAGCCATGCAAAACCTTCCGAATGTTGTTCTTTGACTCGGA 2760
|||||
Db 2701 TGGGAAATGTCTACAGAAAGCCATGCAAAACCTTCCGAATGTTGTTCTTTGACTCGGA 2760
OY 2761 GCAGAGTTTGAATACATGTATGACAGAGATTGCAAGTACCATACAGAACTTTGAG 2820
|||||
Db 2761 GCAGAGTTTGAATACATGTATGACAGAGATTGCAAGTACCATACAGAACTTTGAG 2820
OY 2821 GAAAGCACACATCTTGGGGCTTCATTTCTATCACCCAGGCAATGATATTTTCTA 2880
|||||
Db 2821 GAAAGCACACATCTTGGGGCTTCATTTCTATCACCCAGGCAATGATATTTTCTA 2880
OY 2881 TGCTGGCTGTTCCGGTTTGGTGGCTACTTGTGGCAATGATTCATCACTTTCACAGA 2940
|||||
Db 2881 TGCTGGCTGTTCCGGTTTGGTGGCTACTTGTGGCAATGATTCATCACTTTCACAGA 2940
OY 2941 TGCTGCTGTTTCCGGTTTGGTGGCTACTTGTGGCAATGATTCATCACTTTCACAGA 3000
|||||
Db 2941 TGCTGCTGTTTCCGGTTTGGTGGCTACTTGTGGCAATGATTCATCACTTTCACAGA 3000
OY 2941 TGCTGCTGTTTCCGGTTTGGTGGCTACTTGTGGCAATGATTCATCACTTTCACAGA 3000
|||||
Db 2941 TGCTGCTGTTTCCGGTTTGGTGGCTACTTGTGGCAATGATTCATCACTTTCACAGA 3000
OY 3001 AATTGCTCCTGACTATGCAAAAGCCAAATATCAGACAGCCAGCTCATCATGATCTTGA 3060
|||||
Db 3001 AATTGCTCCTGACTATGCAAAAGCCAAATATCAGACAGCCAGCTCATCATGATCTTGA 3060
OY 3061 AAAAGCCCTGTGATTTGACAGCTACAGCCCTCAGGGCTCAAGCCAAATACCTTGAAG 3120
|||||
Db 3061 AAAAGCCCTGTGATTTGACAGCTACAGCCCTCAGGGCTCAAGCCAAATACCTTGAAG 3120
OY 3121 AAATGTACATTTAATGAGTGTGTTCAACTATCCACTGACAGACATCCCGTGT 3180
|||||
Db 3121 AAATGTACATTTAATGAGTGTGTTCAACTATCCACTGACAGACATCCCGTGT 3180
OY 3181 CCAAGGGCTGAGGCTGAGGTGAAGAAGGCCAGAGCCGTGGCCCTCGAAGGTAGCAGTGG 3240
|||||
Db 3181 CCAAGGGCTGAGGCTGAGGTGAAGAAGGCCAGAGCCGTGGCCCTCGAAGGTAGCAGTGG 3240
OY 3241 CTGTGGGAAGAAGCAAGTTGTTAGCTCTAGAGGCTTCTATGACCCCTTGGCTGCTC 3300
|||||
Db 3241 CTGTGGGAAGAAGCAAGTTGTTAGCTCTAGAGGCTTCTATGACCCCTTGGCTGCTC 3300
OY 3241 CTGTGGGAAGAAGCAAGTTGTTAGCTCTAGAGGCTTCTATGACCCCTTGGCTGCTC 3300
|||||
Db 3241 CTGTGGGAAGAAGCAAGTTGTTAGCTCTAGAGGCTTCTATGACCCCTTGGCTGCTC 3300
OY 3301 AGTGCTAATGTAGTGAAGATTAAGACCTGGAATGTCCAGTGGCTCCGAGCACACT 3360
|||||
Db 3301 AGTGCTAATGTAGTGAAGATTAAGACCTGGAATGTCCAGTGGCTCCGAGCACACT 3360
OY 3361 GGGCATGTGTCTCAGAGGCCATCTGTTGACTGACAGATTCGCGAAMATTTGCTTA 3420
|||||
Db 3361 GGGCATGTGTCTCAGAGGCCATCTGTTGACTGACAGATTCGCGAAMATTTGCTTA 3420
OY 3421 TGGAGACAACAGCGGGCTGTATACATGAAGATTAAGCAAGGACAGGCAAGAGGCCAA 3480
|||||
Db 3421 TGGAGACAACAGCGGGCTGTATACATGAAGATTAAGCAAGGACAGGCAAGAGGCCAA 3480
OY 3481 CATACACCACTTCATGAGACATCCCGAGAAATFACAAACCAAGATAGAACAAAG 3540
|||||
Db 3481 CATACACCACTTCATGAGACATCCCGAGAAATFACAAACCAAGATAGAACAAAG 3540
OY 3541 AACCCAGCTCTGCTGAGGAGCAAGAAACAGGCAATTCAGTCTGGCTTGTGTGACA 3600
|||||

Db 3541 AACCCAGCTCTGTGTGCGCAGAAACAGCGCATTTGCCATAGCTGCGGCTCTTTGTAGAA 3600
Qy 3601 GCTCATATATTTTGTGTTGATGAGTACATAGCTCTGATACGAAAGTGAAGT 3660
Db 3601 GCTCATATATTTTGTGTTGATGAGTACATAGCTCTGATACGAAAGTGAAGT 3660
Qy 3661 TGTCCAAAGAACCCGTGACAAAGCCAGAGAACCCGACCTGATGTGATGCCACCG 3720
Db 3661 TGTCCAAAGAACCCGTGACAAAGCCAGAGAACCCGACCTGATGTGATGCCACCG 3720
Qy 3721 CTGTCCACATCCAGAAATGCAATTTAATAGTGGTGTTCGAATGGCAAGTCAAGA 3780
Db 3721 CTGTCCACATCCAGAAATGCAATTTAATAGTGGTGTTCGAATGGCAAGTCAAGA 3780
Qy 3781 GCATGCGACACATCAACAGCTGCTGCGCAGAAAGCATCTATTTCATGCTGACGT 3840
Db 3781 GCATGCGACACATCAACAGCTGCTGCGCAGAAAGCATCTATTTCATGCTGACGT 3840
Qy 3841 CCAGGCTGGAGCAAGCGCTAGTGAACCTGTGCGCATATGAGCTGTTAAATTTTTTAA 3900
Db 3841 CCAGGCTGGAGCAAGCGCTAGTGAACCTGTGCGCATATGAGCTGTTAAATTTTTTAA 3900
Qy 3901 ATTTGTGTTAAACATGCAATTTAATCAAAAGTTAAAGTGACACTTACTGAAAACT 3960
Db 3901 ATTTGTGTTAAACATGCAATTTAATCAAAAGTTAAAGTGACACTTACTGAAAACT 3960
Qy 3961 ATGTAACTACTGCTTTTACATTTCTGCTGCAACTGAAAGATTCACCAAGTTCAG 4020
Db 3961 ATGTAACTACTGCTTTTACATTTCTGCTGCAACTGAAAGATTCACCAAGTTCAG 4020
Qy 4021 AGCTCTCAGATTTTAAATTTAAAGAACCAAAAGAAACATTATGATGCAATTAATAC 4080
Db 4021 AGCTCTCAGATTTTAAATTTAAAGAACCAAAAGAAACATTATGATGCAATTAATAC 4080
Qy 4081 TGGTGTATTTGCAATTAATTAAGAGTAATTCAAAGTAGATTTTGTAAATTAATG 4140
Db 4081 TGGTGTATTTGCAATTAATTAAGAGTAATTCAAAGTAGATTTTGTAAATTAATG 4140
Qy 4141 TATAATTTTGTATATTTTAAATTTGTAATCTGCTTGTGCAAGATTAATGAAGT 4200
Db 4141 TATAATTTTGTATATTTTAAATTTGTAATCTGCTTGTGCAAGATTAATGAAGT 4200
Qy 4201 GTAAAGACTGTAATGTTGAATTAAGTCTGATTAATTAAGTAACTTTATATC 4260
Db 4201 GTAAAGACTGTAATGTTGAATTAAGTCTGATTAATTAAGTAACTTTATATC 4260
Qy 4261 AAAAAAAAAAAAAAAAAA 4279
Db 4261 AAAAAAAAAAAAAAAAAA 4279

RESULT 5
AF045016 4317 bp mRNA linear MAM 07-FEB-1998
LOCUS Canis familiaris multidrug resistance p-glycoprotein (MDR1) mRNA,
DEFINITION complete cds.
ACCESSION AF045016
VERSION AF045016.1 GI:2852440
KEYWORDS
SOURCE .
ORGANISM Canis familiaris.
Canis familiaris.
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Carnivora; Fissipedia; Canidae; Canis.
REFERENCE
1 (bases 1 to 4317)
Puel,O., Lepage,J.F., Alvinerie,M., Gallier,P. and Pineau,T.
Direct Submission
Submitted (28-JAN-1998) Pharmacology, INRA, BP 3, 180 Chemin de
Tournefeuille, Toulouse Cedex 9 31931, France

FEATURES
source
1..4317
location/Qualifiers
/organism="Canis familiaris"
/db_xref="taxon:9615"
/tissue_type="colon"

gene
CCS
/note="beagle"
1..4317
/gene="MDR1"
70..3912
/gene="MDR1"
/codon_start=1
/product="multidrug resistance p-glycoprotein"
/protein_id="AAC02113.1"
/db_xref="GI:2852441"
/translation="MDPEGKGSLEKFNFMGRKSKREKKKPTVSTPAMPFSNM
LDRLMLGTMALIHGAALPLMLVFNQMDSFANAGISHNKTFPIVINESTNNO
IFINLHEEMTYAYVYSGIGAVLVAAIYOVSPMCAGAOILKIKOPHAIIMOE
IGMPVHVGDELNETLDDYSKINGEDGKMGFOSIATPEFGVITPCKRLIV
ILAIIPVLGSAIWAKTLSFTPEKILLAAKAAVEVLAATRYATVSGKKKEIE
RYNKLLEAKRIGIKKATTANISGAAPFLITYASYALAEWYGSVLSSYTTGOVLT
VFEVSLIGAPSIIGASPSITEFANARAAAEIETIDNKPISIDYSKSGIKPDMING
NYCIDGODIRITINVRHLEITGVVSOEVLFAATTIAENIRYGRNNTMDEIKAVEA
NAYDEIMKLPKFDLIVGRCARLSGOKORIALARLVNRPILKIDEXTSALDRES
EAVVOVALDKARKGRITVIAHRLSTVRNADVIAGPDGVIYERKLGHDLMKEGITYE
KLVMQTRGNEIELENATGESSESDALEMSPDSCSLIKRSTRSTIHAPOGDRK
LGTRKEDLNEVPSVEMRLIKINSTEMPRFVVCIPCAIINGLOPAFSTIFSRIGIF
TRDDPEPKRONSMPEVPLVLCISFTIFEQFPKAGSILIKRLRYWFSML
RODYSWDDPRNTGALTTLRLANDAAVKALISRLAVITONIANIGTIIISLYGW
OLTLIAIPLIPIIAGVEMKMLSGOALKDKRELGAAGKATEALENFRVLSLRE
QKEFYTAOSLOVPEYRNSLRKAHIFGVSFIQTQMMYFVSAGCFREGALVANEPMF
QDLVLSAIVFEGMAVGOVSPADYAKKVAHVIMIEKSPIDSYSPHGKPN
TLBGNVFNENEVNYPRPDIPVLOGISLEVKKGOTLAIYVSGSGCKSTVOLLREY
DPLAGSVLIDGKEIKHLNVOMLRAHIGIYSOEPILIDCSIAEMIAIGDSRVLSHEEY
MOAKKEANIHLEITLPEKYNTRVSGKTIOLSGOKORIALARLVNRPILKIDEXAT
SALDRESKYVOALDKARKGRITVIAHRLSTIONADLIVFONGKVEHGTQOOLL
AQKRTYSMSVAGAR"

BASE COUNT 1293 a 844 c 1019 g 1161 t

ORIGIN

Query Match 98.7%; Score 4221.8; DB 4; Length 4317;
Best Local Similarity 99.5%; Pred. No. 0;
Matches 4245; Conservative 0; Mismatches 17; Indels 3; Gaps 1;

Qy 3 AGCCGAGAGTGGGATGATCTGTAAGAGCGCGTAAGGGAGTGCAGAGAATCTCT 62
Db 56 AGCCGAGAGTGGGATGATCTGTAAGAGCGCGTAAGGGAGTGCAGAGAATCTCT 115
Qy 63 GGAATAATGGGCAAAAAAGTAAAAAATGAGAAAGAAAGAAAGAAACCACTGTACGA 122
Db 116 GGAATAATGGGCAAAAAAGTAAAAA--GAGAAAGAAAGAAAGAAACCACTGTACGA 172
Qy 123 CGTTTGCATGTTTGGCTATTCAAATTTGGCTGATAGTTGATATGTTGGGGACAA 182
Db 173 CGTTTGCATGTTTGGCTATTCAAATTTGGCTGATAGTTGATATGTTGGGGACAA 232
Qy 183 TGGCTGCATCATCCATGAGAGTGCACCTCCCTCTCATGATGCTGTTTGGAAACATGA 242
Db 233 TGGCTGCATCATCCATGAGAGTGCACCTCCCTCTCATGATGCTGTTTGGAAACATGA 292
Qy 243 CAGATAGCTTTGCAAAATGCAAGAAATTTCAAGAAACAAACTTTTCCATTAATTAATG 302
Db 293 CAGATAGCTTTGCAAAATGCAAGAAATTTCAAGAAACAAACTTTTCCATTAATTAATG 352
Qy 303 AAGATTAATGGAACAATACACAACATTTTCATCAACCATCTGGAGAGAAATGACACGT 362
Db 353 AAGATTAATGGAACAATACACAACATTTTCATCAACCATCTGGAGAGAAATGACACGT 412
Qy 363 ATGCCATATTATTAAGTGGATGCTGCTGGGGGTCGCTGGTGGCTTACATCCAGGTTT 422
Db 413 ATGCCATATTATTAAGTGGATGCTGCTGGGGGTCGCTGGTGGCTTACATCCAGGTTT 472
Qy 423 CATTTGCTGCTGGCAGACAGACAGATACATCAAAATTAAGAAACAAATTTTTCATG 482
Db 473 CATTTGCTGCTGGCAGACAGACAGATACATCAAAATTAAGAAACAAATTTTTCATG 532
Qy 483 CATCATCGACAGAGATTTGGCTGTTGACGTGCATGACGTTGGGAGCTTAACACC 542

Db 533 CTATCATGGACAGAGATTGGCTGGTTGACGTGCATGACGTTGGGAGCTTAACACC 592
Qy 543 GGCTCACAGACGATGTCCTCCAAATCAATGAGGAATTTGGCGACAAATTTGAAATGTTCT 602
Db 593 GGCTCACAGACGATGTCCTCCAAATCAATGAGGAATTTGGCGACAAATTTGAAATGTTCT 652
Qy 603 TTCACTCAATAGCAACATTTTTCACCGGTTTATAGTGGGGTTTACAGTGGTTTGAAGC 662
Db 653 TTCAATCAATAGCAACATTTTTCACCGGTTTATAGTGGGGTTTACAGTGGTTTGAAGC 712
Qy 663 TAACCCCTGGATTGGGACCATCAGCCCTGTTCTTGGACTTTCAGCCGCCATCTGGGCA 722
Db 713 TAACCCCTGGATTGGGACCATCAGCCCTGTTCTTGGACTTTCAGCCGCCATCTGGGCA 772
Qy 723 AGATACATCTTCATTTACTGATPAAGAACTCTTGGCCTATGCAAAAGTGGAGCACTAG 782
Db 773 AGATACATCTTCATTTACTGATPAAGAACTCTTGGCCTATGCAAAAGTGGAGCACTAG 832
Qy 783 CTGAAGAAGTCTTAGCAGCAATCAGAACTGATGTCCTTGGAGGACAAAGAAAGAAC 842
Db 833 CTGAAGAAGTCTTAGCAGCAATCAGAACTGATGTCCTTGGAGGACAAAGAAAGAAC 892
Qy 843 TTGAAGGTACACAAAAAATTTAGAGAAGCTAAAGGAATGGGATPAAGAAAGCTATCA 902
Db 893 TTGAAGGTACACAAAAAATTTAGAGAAGCTAAAGGAATGGGATPAAGAAAGCTATCA 952
Qy 903 CGGCCACATTTCTATTGGTCCCTCTTCTATGATCTATGATCTATGCTCTGCTGCTT 962
Db 953 CGGCCACATTTCTATTGGTCCCTCTTCTATGATCTATGATCTATGCTCTGCTGCTT 1012
Qy 963 TCTGGTATGGGACCTCTTGGTCTCTCCATGCAATATTTCTATTGGAAGTACTACTG 1022
Db 1013 TCTGGTATGGGACCTCTTGGTCTCTCCATGCAATATTTCTATTGGAAGTACTACTG 1072
Qy 1023 TCTTCTTTTCTGATTAATTTGGGGCTTTTACTATTGGACAGATGCCCAAGCTTAAG 1082
Db 1073 TCTTCTTTTCTGATTAATTTGGGGCTTTTACTATTGGACAGATGCCCAAGCTTAAG 1132
Qy 1083 CATTTGCAAGCAGAGAGAGAGCTTATGAAATCTTCAAGATATTTGACATTAACCA 1142
Db 1133 CATTTGCAAGCAGAGAGAGAGCTTATGAAATCTTCAAGATATTTGACATTAACCA 1192
Qy 1143 GCATTCAGAGCTATTCGGAAGATGAGCATTAACAGATATTTAAGGAAATTTGCAAT 1202
Db 1193 GCATTCAGAGCTATTCGGAAGATGAGCATTAACAGATATTTAAGGAAATTTGCAAT 1252
Qy 1203 TCAAAAATGTTCACTTCACTTACCCTTCTCGAAAAGAAATTAAGATCTTAAAGGCTCA 1262
Db 1253 TCAAAAATGTTCACTTCACTTACCCTTCTCGAAAAGAAATTAAGATCTTAAAGGCTCA 1312
Qy 1263 ACCTGAAGGTTCAAGAGTGGGACAGATGCGCTGGTTGGGAACAGTGGCTCGGGAGA 1322
Db 1313 ACCTGAAGGTTCAAGAGTGGGACAGATGCGCTGGTTGGGAACAGTGGCTCGGGAGA 1372
Qy 1323 GCAGAGACCGTGCAGCTGATCAGAGGCTATGACCCACAGATGGCATGCTGTATTTG 1382
Db 1373 GCAGAGACCGTGCAGCTGATCAGAGGCTATGACCCACAGATGGCATGCTGTATTTG 1432
Qy 1383 ATGACAGAGACATTAGGACCATTAATGTAAGCATCTTGGGAAATTTACTGGTGGTGTA 1442
Db 1433 ATGACAGAGACATTAGGACCATTAATGTAAGCATCTTGGGAAATTTACTGGTGGTGTA 1492
Qy 1443 GTCAGAGACCTGTGTGTTGGCCACCGATAGCTGAAAACATTTGCCCTATGGCCGGAAA 1502
Db 1493 GTCAGAGACCTGTGTGTTGGCCACCGATAGCTGAAAACATTTGCCCTATGGCCGGAAA 1552
Qy 1503 ATGACACCATGATGAGATTGAGAAAGCTGTTAAGGAAGCAATGCCATGATTTTATCA 1562
Db 1553 ATGACACCATGATGAGATTGAGAAAGCTGTTAAGGAAGCAATGCCATGATTTTATCA 1612
Qy 1563 TGAACCTACCTAATTAATTTGACACTGTGGTTGGAGAGAGAGGGCCCGGCTAGTGGTG 1622
Db 1613 TGAACCTACCTAATTAATTTGACACTGTGGTTGGAGAGAGAGGGCCCGGCTAGTGGTG 1672

Qy 1623 GACAGAAACAGAGAAATCCGCATTTGCTCGGGCCCTGGTTCCGACCCCAAGATTTCTTGC 1682
Db 1673 GACAGAAACAGAGAAATCCGCATTTGCTCGGGCCCTGGTTCCGACCCCAAGATTTCTTGC 1732
Qy 1683 TGGATGAGCAACGTCAGCTTGCACACTGAAAGTGAAGCAGTGGTTCAGTGGCCCTGG 1742
Db 1733 TGGATGAGCAACGTCAGCTTGCAGACTGAAAGTGAAGCAGTGGTTCAGTGGCCCTGG 1792
Qy 1743 ATAGGCCAAGAAAGGCCGAGACTACCATTTGTATAGTCATCGTTTGTCTACAGTGGTA 1802
Db 1793 ATAGGCCAAGAAAGGCCGAGACTACCATTTGTATAGTCATCGTTTGTCTACAGTGGTA 1852
Qy 1803 ATGCCCATGTCATTTGGTGGTTGATGATGAGTCAATTTGAGAGAAATCATGATG 1862
Db 1853 ATGCCCATGTCATTTGGTGGTTGATGATGAGTCAATTTGAGAGAAATCATGATG 1912
Qy 1863 AACCTATGAAGAAGAGAGGCGCATTTACTTCAAACTTGTCCAAATGCAAGCAAGAGAAATG 1922
Db 1913 AACCTATGAAGAAGAGAGGCGCATTTACTTCAAACTTGTCCAAATGCAAGCAAGAGAAATG 1972
Qy 1923 AAATGAGTTAGAAAATGCCACGTTGGAATCCAAAGTGAAGTGAATGCTTGAATGT 1982
Db 1973 AAATGAGTTAGAAAATGCCACGTTGGAATCCAAAGTGAAGTGAATGCTTGAATGT 2032
Qy 1983 CTCCAAAAGATTCAAGGCTCCAGTTTAATAAAGAAAGATCACTCCAGAGATATACATG 2042
Db 2033 CTCCAAAAGATTCAAGGCTCCAGTTTAATAAAGAAAGATCACTCCAGAGATATACATG 2092
Qy 2043 CACCAAGAGCCCAAGACAGAAAGCTTGTACAAAAGAGGACTTGAATGAAATGTACTCT 2102
Db 2093 CACCAAGAGCCCAAGACAGAAAGCTTGTACAAAAGAGGACTTGAATGAAATGTACTCT 2152
Qy 2103 CAGTTTCTTCTGGAGATCTGAAAGTGAATCACTCAATGATGAGCTTATTTGTGCTTG 2162
Db 2153 CAGTTTCTTCTGGAGATCTGAAAGTGAATCACTCAATGATGAGCTTATTTGTGCTTG 2212
Qy 2163 GTATATTTTGTCTATTTATTAACGAGGCTTGCACACAGATTTTCAATATATTTTCA 2222
Db 2213 GTATATTTTGTCTATTTATTAACGAGGCTTGCACACAGATTTTCAATATATTTTCA 2272
Qy 2223 GGATTAATGAGATCTTTACCGGAGATGAGATCTGAAAACAAAACGAGATATGATACA 2282
Db 2273 GGATTAATGAGATCTTTACCGGAGATGAGATCTGAAAACAAAACGAGATATGATACA 2332
Qy 2283 TGTTTTCTGATTTGTTTCTAGTCTTGGAAATTTTCTTATTAACATTTTCTCCAGG 2342
Db 2333 TGTTTTCTGATTTGTTTCTAGTCTTGGAAATTTTCTTATTAACATTTTCTCCAGG 2392
Qy 2343 GCTTCACATTTGGCAAGGCTGGGAGATCTCATAAGCGGCTTCGATACATGCTTTCA 2402
Db 2393 GCTTCACATTTGGCAAGGCTGGGAGATCTCATAAGCGGCTTCGATACATGCTTTCA 2452
Qy 2403 GATCCATGCTGAGACAGAGATGTCAGTGTGTTGATGACCTTAATAAACCCAGCTGAGCAT 2462
Db 2453 GATCCATGCTGAGACAGAGATGTCAGTGTGTTGATGACCTTAATAAACCCAGCTGAGCAT 2512
Qy 2463 TGACAAACAGGCTTGGCAATGATGCGCTCAAGTTAAAGGGGCTTATGATTCAGAGCTTG 2522
Db 2513 TGACAAACAGGCTTGGCAATGATGCGCTCAAGTTAAAGGGGCTTATGATTCAGAGCTTG 2572
Qy 2523 CTGTCATTTACCCAGAAATATGCAAAATCTTGGGACAGGCTATTTATTCCTTAATCATG 2582
Db 2573 CTGTCATTTACCCAGAAATATGCAAAATCTTGGGACAGGCTATTTATTCCTTAATCATG 2632
Qy 2583 GTTGGCAATTAACACTTTTACTCTTAGCAATTTGATACCATATTCGAAATAGCAGAGTTG 2642
Db 2633 GTTGGCAATTAACACTTTTACTCTTAGCAATTTGATACCATATTCGAAATAGCAGAGTTG 2692
Qy 2643 TTGAAATGAAGAAATGTTGTCGACAAAGCACTGAAAGATTAAGAAAGAGTGAAGAGCTG 2702
Db 2693 TTGAAATGAAGAAATGTTGTCGACAAAGCACTGAAAGATTAAGAAAGAGTGAAGAGCTG 2752

QY 2703 GGAAAGTTGCTACAGAACCATCGAAAACTTCGGAAGTGTGTTCTTGACTCGGGAGC 2762
|||||
Db 2753 GGAAAGTTGCTACAGAACCATCGAAAACTTCGGAAGTGTGTTCTTGACTCGGGAGC 2812
|||||
QY 2763 AGAAGTTGAATACATGTAATGACAGAGATTGGCAAGTACCATACAGAAACTCTTTGAGGA 2822
|||||
Db 2813 AGAAGTTGAATACATGTAATGACAGAGATTGGCAAGTACCATACAGAAACTCTTTGAGGA 2872
|||||
QY 2823 AAGCACATCTTGGGGTCTCATTTTCTATCACCAGGCAAGTATGTTTCTATG 2882
|||||
Db 2873 AAGCACATCTTGGGGTCTCATTTTCTATCACCAGGCAAGTATGTTTCTATG 2932
|||||
QY 2883 CTGGCTGTTCCGGTTTGTTGGTCCCTACCTTGTTGGCAAAATGAGTTTCAGACTTTCAGAGT 2942
2933 CTGGCTGTTCCGGTTTGTTGGTCCCTACCTTGTTGGCAAAATGAGTTTCAGACTTTCAGAGT 2992
QY 2943 TTCTTTTGTATGTTCTACCTATGCTTGTGGTCCATGGAGAGTGGGGCAGTCACTTTCAT 3002
|||||
Db 2993 TTCTTTTGTATGTTCTACCTATGCTTGTGGTCCATGGAGAGTGGGGCAGTCACTTTCAT 3052
|||||
QY 3003 TTGCTCTGACTATGCAAAAGCCAAAGTATCAGAGCCACGTCATCATGATCATTTGAAA 3062
Db 3053 TTGCTCTGACTATGCAAAAGCCAAAGTATCAGAGCCACGTCATCATGATCATTTGAAA 3112
QY 3063 AAAGCCCTGATTTGACAGCTACAGCCCTCAAGCCCTCAAGCCAAATACGTTTGAAGGAA 3122
Db 3113 AAAGCCCTGATTTGACAGCTACAGCCCTCAAGCCCTCAAGCCAAATACGTTTGAAGGAA 3172
QY 3123 ATGTGACATTTAATGAGAGTGTGTTCAACTATCCATCGACAGACATCCCGTGTCTCC 3182
Db 3173 ATGTGACATTTAATGAGAGTGTGTTCAACTATCCATCGACAGACATCCCGTGTCTCC 3232
QY 3183 AGGGGCTGAGCTCTGAGGTGAAGAGGGCCAGAGCGTGGCCCTCTGATGATGAGTGTCT 3242
Db 3233 AGGGGCTGAGCTCTGAGGTGAAGAGGGCCAGAGCGTGGCCCTCTGATGATGAGTGTCT 3292
QY 3243 GTGGGAAGAGACAGTGTGTTCACTCTCTAGAGCGCTTCTATGACCCCTTGCTGTTGACG 3302
Db 3293 GTGGGAAGAGACAGTGTGTTCACTCTCTAGAGCGCTTCTATGACCCCTTGCTGTTGACG 3352
QY 3303 TGCTAATTTGATGGCAAAAGATTAAGACACCTGTAATGTCAGTGGGCTCCAGACACCTGG 3362
Db 3353 TGCTAATTTGATGGCAAAAGATTAAGACACCTGTAATGTCAGTGGGCTCCAGACACCTGG 3412
QY 3363 GCATCGTGTCTCAGAGCCCATCTGTTTGTACTGACGACATTCGCCGAACAATTTGCTTATG 3422
Db 3413 GCATCGTGTCTCAGAGCCCATCTGTTTGTACTGACGACATTCGCCGAACAATTTGCTTATG 3472
QY 3423 GAGACAACAGCCGGGTGATCATCATGAGAGATTAATGAGAGCAGCCAGAGGCCAACA 3482
Db 3473 GAGACAACAGCCGGGTGATCATCATGAGAGATTAATGAGAGCAGCCAGAGGCCAACA 3532
QY 3483 TACACCACTTCATCGAGACATCCCTGAGAAATATACACCCAGAGTATAGAGCAAGGAA 3542
Db 3533 TACACCACTTCATCGAGACATCCCTGAGAAATATACACCCAGAGTATAGAGCAAGGAA 3592
QY 3543 CCAGGCTCTGTTGGTGGCCGAAAGAGCGATTTGCCATAGCTCGGCTCTTTGTTAGACAGC 3602
Db 3593 CCAGGCTCTGTTGGTGGCCGAAAGAGCGATTTGCCATAGCTCGGCTCTTTGTTAGACAGC 3652
QY 3603 CTCATATTTTGGTGTGATGTAAGTATCATCGTCTGATATAGAAAGTGAAGAAAGGTG 3662
Db 3653 CTCATATTTTGGTGTGATGTAAGTATCATCGTCTGATATAGAAAGTGAAGAAAGGTG 3712
QY 3663 TCCAAAGACCCCTGAGCAAGCCAGAGAGGCGCAGCTGCAATGTGATCGCCACCGCT 3722
Db 3713 TCCAAAGACCCCTGAGCAAGCCAGAGAGGCGCAGCTGCAATGTGATCGCCACCGCT 3772
QY 3723 TGTCCACCATTCAGAAATGAGATTAATAGTGTGTTCAAAATGCGAAAGTCAAGAGC 3782
Db 3773 TGTCCACCATTCAGAAATGAGATTAATAGTGTGTTCAAAATGCGAAAGTCAAGAGC 3832
QY 3783 ATGGCACACATCAACAGCTGCTGGCCGAAAGAGCATCTATTTTTCATGATGTCAGTCTCC 3842

|||||
Db 3833 ATGGCACATCAACAGCTGCTGGCTCGAAAGCATCTATTTTCCATGATCAGTGTCC 3892
QY 3843 AGGCTGAGCAAAAGCGCTAGTGAAGTGTGGCCATATGAGCTTAATATTTTATATAT 3902
Db 3893 AGGCTGAGCAAAAGCGCTAGTGAAGTGTGGCCATATGAGCTTAATATTTTATATAT 3952
QY 3903 TTGTGTTAAACATGCGCATTTTAATCAAAAGTTAAAGGTGAGCATTTACTGGAAAACTAT 3962
Db 3953 TTGTGTTAAACATGCGCATTTTAATCAAAAGTTAAAGGTGAGCATTTACTGGAAAACTAT 4012
QY 3963 GTAGAACTACCTGTTTAACTTTCTGCTGCAACTGAAGATTCATTCACCAAGTTAGAG 4022
Db 4013 GTAGAACTACCTGTTTAACTTTCTGCTGCAACTGAAGATTCATTCACCAAGTTAGAG 4072
QY 4023 TCTTCAGATTTTAAATTAAGAAACCAAAACATTCATGATGAGTAAATATGACG 4082
Db 4073 TCTTCAGATTTTAAATTAAGAAACCAAAACATTCATGATGAGTAAATATGACG 4132
QY 4083 GTGTTAATTCATTAATTAATATATAGAGTAATTCAAAGTATGATTTGTTAATTAATGTA 4142
Db 4133 GTGTTAATTCATTAATTAATATATAGAGTAATTCAAAGTATGATTTGTTAATTAATGTA 4192
QY 4143 TAATTTTGTATATATTTTATTTTGTACTACTACTGCTTGTGCGAAGATATAGAACTGT 4202
Db 4193 TAATTTTGTATATTTTATTTTGTACTACTACTGCTTGTGCGAAGATATAGAACTGT 4252
QY 4203 AAAAAGTACTGAATGTTTGAATTAAGTGTAGCTATATTAATTAACCTTTTATATCA 4262
Db 4253 AAAAAGTACTGAATGTTTGAATTAAGTGTAGCTATATTAATTAACCTTTTATATCA 4312
QY 4263 AAAAA 4267
Db 4313 AAAAA 4317
RESULT 6
AX105059
LOCUS
DEFINITION
AX105059
VERSION
AX105059.1 GI:13921211
KEYWORDS
SOURCE
dog.
ORGANISM
Canis familiaris
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Carnivora; Fissipedia; Canidae; Canis.
REFERENCE
1 (bases 1 to 4317)
Stocker,P.J., Steinel-Crespi,D.T., Crespi,C.L., Reif,T.C. and
patten,C.J.
TITLE
P-glycoproteins and uses thereof
JOURNAL
Patent: WO 0123540-A 3 05-Apr-2001;
GENTEST CORPORATION (US)
FEATURES
source
1..4317
location/Qualifiers
/organism="Canis familiaris"
/db_xref="taxon:9615"
70..3912
/note="unnamed protein product"
/codon_start=1
/protein_id="CAC37719.1"
/db_xref="GI:13921212"
/translation="MDPEGGRKGSAEKNFPMKGRKSKKEKKKPTVSTPFAFRISNW
LDRLVLMVGTMAATIRHGAALPIMLVGNNTDSEFANAGISRNKTFPVITINESTNTO
HEFHLEBMTLYYYISGIGAVLVAATIQVSEFWCLAGROILKIRKOFPAHMOE
IGWHDHVGELNTRLTDVSKINEGIDGVNMFQSIATFFGFTVGTPEMKLTLV
ILAI5PVLGSAALIAWALISFDKELIAYAKAGAVEYLAARIVYAFGOKKLE
RYNNLEBARKIGIKAITANISIGAAFLIYASVALAPWYGSILVSEYITIGVIT
VFEFVLIGAFSIGOASPISEFAPNARGAAVEIRKIIDNPSIDYSKSGHRDNDIKN
LEFNHVESYPSRREYKILKGLMKYOSQGTVALVNSGCSSTYQLOMRILYPTDG
MVCIDGODITIVRHRRETTGVYSEDPVLPATTTAENTRGRENTYDEIKAVKA
NAYDFIKRLNRKEDTLVGERARLSGOKORAIATVALRNKRIILLDEATSDLTES
EAVYVALDKRRGRITTVIAHRLSTVRNADVIAGPDGVIVIEKGNHDELMKEGIYE

Db	833	CTGAAAGAGCTTTAGCAGCAATCAGAACTGTGATGTCCCTTTGGAGACGACAAAAGAAAGAAC	892
Qy	843	TTGAAAGCTACAAACAAAATTTAGAGAAGCTAAAGAAATTTGGGATTAAGAAAGCTATCA	902
Db	893	TTGAAAGGTACAAACAAAATTTAGAGAAGCTAAAGAAATTTGGGATTAAGAAAGCTATCA	952
Qy	903	CGGCGCAACATTTCTATTTGTGTGCCCTTTCTTATTTGATCTATAGCATCATATATGCTGTGCTT	962
Db	953	CGGCGCAACATTTCTATTTGTGTGCCCTTTCTTATTTGATCTATAGCATCATATATGCTGTGCTT	1012
Qy	963	TTCTGGTATGGGACCTCTGTGGCTCTCTCAGAGAAATATTTCTATTTGGACAGTACTCAGTG	1022
Db	1013	TTCTGGTATGGGACCTCTGTGGCTCTCTCAGAGAAATATTTCTATTTGGACAGTACTCAGTG	1072
Qy	1023	TCCTCTTTTCTGTATTAATTTGGGGCTTTTATGATTTGGACAGGCACTCCCAAGATTTGAAG	1082
Db	1073	TCCTCTTTTCTGTATTAATTTGGGGCTTTTATGATTTGGACAGGCACTCCCAAGATTTGAAG	1132
Qy	1083	CATTTGCAAAACGCAAGAGAGCGAGCTTATGAATCTTCAATATTAATTAAGCAATTAACCCA	1142
Db	1133	CATTTGCAAAACGCAAGAGAGCGAGCTTATGAATCTTCAATATTAATTAAGCAATTAACCCA	1192
Qy	1143	GCATTTGACAGCTATTCGAAAGATGTGACATTAACAGATTAATTTAAGGAAATTTGGAAAT	1202
Db	1193	GCATTTGACAGCTATTCGAAAGATGTGACATTAACAGATTAATTTAAGGAAATTTGGAAAT	1252
Qy	1203	TCAAAAATGTTCACTTCAGTTACCTTCCTTCGAAAAGAAAGTTAATAACTTAAGAGGCTCTCA	1262
Db	1253	TCAAAAATGTTCACTTCAGTTACCTTCCTTCGAAAAGAAAGTTAATAACTTAAGAGGCTCTCA	1312
Qy	1263	ACCTGAAGCTTCAGAGTGGGCGACAGTGGCCCTGGTTGGGAACAGTGGCTGCGCGGAAGA	1322
Db	1313	ACCTGAAGCTTCAGAGTGGGCGACAGTGGCCCTGGTTGGGAACAGTGGCTGCGCGGAAGA	1372
Qy	1323	GCACAGACCGTCAGCTGATGCAGAGGCTCTATGACCCACAGATGTGCTGTCTGTGTTG	1382
Db	1373	GCACAGACCGTCAGCTGATGCAGAGGCTCTATGACCCACAGATGTGCTGTCTGTGTTG	1432
Qy	1383	ATGGACAGAGACATTTAGACACCTAAATGTAAGGATCTTGGGAAATTTACTGCTGTGATGA	1442
Db	1433	ATGGACAGAGACATTTAGACACCTAAATGTAAGGATCTTGGGAAATTTACTGCTGTGATGA	1492
Qy	1443	GTCAGAGAGCTGTGTTGTTTGGCCACACGATAGCTGAAAACATTTGCTATGGCCGCGAAA	1502
Db	1493	GTCAGAGAGCTGTGTTGTTTGGCCACACGATAGCTGAAAACATTTGCTATGGCCGCGAAA	1552
Qy	1503	ATGTCCACATGTGATGAGATTGAGAAAGCTTTAAGAAACCAATGCCCTATATTTTATNCA	1562
Db	1553	ATGTCCACATGTGATGAGATTGAGAAAGCTTTAAGAAACCAATGCCCTATATTTTATNCA	1612
Qy	1563	TGAAACTACTTAATTAATTTGACACTCTGTGTTGGAGAGAGAGGGGCCAGCTGTAGTGTG	1622
Db	1613	TGAAACTACTTAATTAATTTGACACTCTGTGTTGGAGAGAGAGGGGCCAGCTGTAGTGTG	1672
Qy	1623	GACAGAAACAGAGAAATCGCAATTCCTCGGGCCCTGTTGGCAACCCCAAGATTCCTTCTGC	1682
Db	1673	GACAGAAACAGAGAAATCGCAATTCCTCGGGCCCTGTTGGCAACCCCAAGATTCCTTCTGC	1732
Qy	1683	TGGATGAGCAACGTCACACTCTGACACTGAAAGTGAAGAGTGTTCAGTGGGCCCTCGG	1742
Db	1733	TGGATGAGCAACGTCACACTCTGACACTGAAAGTGAAGAGTGTTCAGTGGGCCCTCGG	1792
Qy	1743	ATTAAGGCCAGAAAAGCCGCGACATACATTTGATAGCTCATCTGTTTGTCTACAGTTGCTTA	1802
Db	1793	ATTAAGGCCAGAAAAGCCGCGACATACATTTGATAGCTCATCTGTTTGTCTACAGTTGCTTA	1852
Qy	1803	ATGCCGATGTCAATGCTGTGTTTGATGATGATGATGATGATGATGATGATGATGATGATGATG	1866
Db	1853	ATGCCGATGTCAATGCTGTGTTTGATGATGATGATGATGATGATGATGATGATGATGATGATG	1912
Qy	1863	AACTCATGAAAGAGAGGCAATTTACTTCAAACTTGGCACAAATGCAAGCAAGAGGAAATG	1922
Db	1913	AACTCATGAAAGAGAGGCAATTTACTTCAAACTTGGCACAAATGCAAGCAAGAGGAAATG	1972

QY 1923 AATTGAGTTAGAAAATGCGACTGCTGTAATCCAAAAGTGAAGTGAATGCTTGGAAATGT 1982
|||||
Db 1973 AATTGAGTTAGAAAATGCGACTGCTGTAATCCAAAAGTGAAGTGAATGCTTGGAAATGT 2032
QY 1983 CTCCAAAAAGATTCAAGGCTCAGTTTAATAAAAAGATCAACCTCCAGAGATATCAATG 2042
|||||
Db 2033 CTCCAAAAAGATTCAAGGCTCAGTTTAATAAAAAGATCAACCTCCAGAGATATCAATG 2092
QY 2043 CACCAACAAGGCCAAGACAGAAAGCTTGGTACAAAAGAGCATTTGAATGGAATGTACCTC 2102
|||||
Db 2093 CACCAACAAGGCCAAGACAGAAAGCTTGGTACAAAAGAGCATTTGAATGGAATGTACCTC 2152
QY 2103 CAGTTTCTCTTGAGAGATTCTGAAGCTGAACCAACTGAATGAGCCCTTTTGTGGTTG 2162
|||||
Db 2153 CAGTTTCTCTTGAGAGATTCTGAAGCTGAACCAACTGAATGAGCCCTTTTGTGGTTG 2212
QY 2163 GTATATTTTGTGCTATTTAATAAGGAGCCCTGCAACCAACTTTTCAATTAATTTTCAA 2222
|||||
Db 2213 GTATATTTTGTGCTATTTAATAAGGAGCCCTGCAACCAACTTTTCAATTAATTTTCAA 2272
QY 2223 GGATTAATAGGGAATCTTACCAGATGAGGATCTGAAAACAAAAGACAGAAATGTAAACA 2282
|||||
Db 2273 GGATTAATAGGGAATCTTACCAGATGAGGATCTGAAAACAAAAGACAGAAATGTAAACA 2332
QY 2283 TGTTTCTGTAATGTTTCTAGTCTTGGATTAATTTCTTTATTAACAATTTTCTCCAGG 2342
|||||
Db 2333 TGTTTCTGTAATGTTTCTAGTCTTGGATTAATTTCTTTTATTACATTTTCTCCAGG 2392
QY 2343 GCTTTCATTTGGCAAAAGCTGGGAGATCTCAGTAAAGGCGCTCGATACATGGTTTCA 2402
|||||
Db 2393 GCTTTCATTTGGCAAAAGCTGGGAGATCTCAGTAAAGGCGCTCGATACATGGTTTCA 2452
QY 2403 GATCCATGCTGAGACAGATGTCAGCTGTTGATGACCCCTAAAAAACCAACGAGAGCAT 2462
|||||
Db 2453 GATCCATGCTGAGACAGATGTCAGCTGTTGATGACCCCTAAAAAACCAACGAGAGCAT 2512
QY 2463 TGACAACAAGGCTTGCATATGATGCGGCTCAAGTTAAAGGCGCTGATAGGTTTCAAGGCTTG 2522
|||||
Db 2513 TGACAACAAGGCTTGCATATGATGCGGCTCAAGTTAAAGGCGCTGATAGGTTTCAAGGCTTG 2572
QY 2523 CTGTATTAACCCAGAAATATACAAATCTTGGACAGGCAATTAATATCTTAACTATG 2582
|||||
Db 2573 CTGTATTAACCCAGAAATATACAAATCTTGGACAGGCAATTAATATCTTAACTATG 2632
QY 2583 GTTGGCAATTAACACTTTTACTCTTACCAATGTACCCCAATGCAATGCAATGAGAGTTG 2642
|||||
Db 2633 GTTGGCAATTAACACTTTTACTCTTACCAATGTACCCCAATGCAATGCAATGAGAGTTG 2692
QY 2643 TTGAAATGAAAATGTGTCTGCGAACAGCTGAAGATGAAGAAGAGCTAGAAAGAGCTG 2702
|||||
Db 2693 TTGAAATGAAAATGTGTCTGCGAACAGCTGAAGATGAAGAAGAGCTAGAAAGAGCTG 2752
QY 2703 GGAAGATTGCTACAGAAGCCATGGAACACTTCCGAACCTGTTTCTTGACTGGGAGC 2762
|||||
Db 2753 GGAAGATTGCTACAGAAGCCATGGAACACTTCCGAACCTGTTTCTTGACTGGGAGC 2812
QY 2763 AGAAGTTGAATCATGTATGCAAGAGTTTSCAAGTACCATACAGAAACCTTTGAGGA 2822
|||||
Db 2813 AGAAGTTGAATCATGTATGCAAGAGTTTSCAAGTACCATACAGAAACCTTTGAGGA 2872
QY 2823 AAGCAGACATCTTGGGGCTCTCATTTTCTATCACCCAGGCAATATGTATTTTCTATG 2882
|||||
Db 2873 AAGCAGACATCTTGGGGCTCTCATTTTCTATCACCCAGGCAATATGTATTTTCTATG 2932
QY 2883 CTGGCTTTTCCGGTTGGTGTCTACTTGTGGCAAAATGAATTCATGAATTTTCTAGATG 2942
|||||
Db 2933 CTGGCTTTTCCGGTTGGTGTCTACTTGTGGCAAAATGAATTCATGAATTTTCTAGATG 2992
QY 2943 TTTCTTTGGTATTTCACTATGTCTTGTGGTGCATGAGAGGGGCGAGCTCAGTTCA 3002
|||||
Db 2993 TTTCTTTGGTATTTCACTATGTCTTGTGGTGCATGAGAGGGGCGAGCTCAGTTCA 3052

QY 3003 TTGCTCTGTACTATGCCAAAAGCCAAAGTATCAGACGCCACGCTCATCATTTGAA 3062
|||||
Db 3053 TTGCTCTGTACTATGCCAAAAGCCAAAGTATCAGACGCCACGCTCATCATTTGAA 3112
QY 3063 AAAGCCCTTGATTTGACACTACACCCCTCAGAGGCTTAAGCCAAATACGTTGGAGAA 3122
|||||
Db 3113 AAAGCCCTTGATTTGACACTACACCCCTCAGAGGCTTAAGCCAAATACGTTGGAGAA 3172
QY 3123 ATGTGACATTTAATGAGTGTGTTCACATATCCACTGCGCCAGACATATCCCTGCTCC 3182
|||||
Db 3173 ATGTGACATTTAATGAGTGTGTTCACATATCCACTGCGCCAGACATATCCCTGCTCC 3232
QY 3183 AAGGAGTGAGCCCTGAGGTGAAGAAGGCGCACAGCTGGCCCTGTAGTAGAGTGGCT 3242
|||||
Db 3233 AAGGAGTGAGCCCTGAGGTGAAGAAGGCGCACAGCTGGCCCTGTAGTAGAGTGGCT 3292
QY 3243 GTGGGAAGACACAGTTTTCAGCTCTAGAGGCTCTTATGACCCCTTGGCTGGTTCAG 3302
|||||
Db 3293 GTGGGAAGACACAGTTTTCAGCTCTAGAGGCTCTTATGACCCCTTGGCTGGTTCAG 3352
QY 3303 TGTCTAATGATGGAAGAGATTAAGCAGCTGAATGTCCAGTGGCTCCGAGCACACTGG 3362
|||||
Db 3353 TGTCTAATGATGGAAGAGATTAAGCAGCTGAATGTCCAGTGGCTCCGAGCACACTGG 3412
QY 3363 GCATCGTGTCTCAGAGACCCCATCTGTTTGAACGACATTCGCCGAGAACTTGGCTATG 3422
|||||
Db 3413 GCATCGTGTCTCAGAGACCCCATCTGTTTGAACGACATTCGCCGAGAACTTGGCTATG 3472
QY 3423 GAGCAACACGCGGCTGTATCATATGAAGAATATAGCAGGACCCCAAGAGGCCAAACA 3482
|||||
Db 3473 GAGCAACACGCGGCTGTATCATATGAAGAATATAGCAGGACCCCAAGAGGCCAAACA 3532
QY 3483 TACACCACTTTCATGAGACACTCCCTGAGAAATACAAACAGAGATAGAGACAAAGAA 3542
|||||
Db 3533 TACACCACTTTCATGAGACACTCCCTGAGAAATACAAACAGAGATAGAGACAAAGAA 3592
QY 3543 CCCAGCTCTCTGGTGCCGAGAAACAGGCAATTTGCCATAGCTGGCTTGTAGACAGC 3602
|||||
Db 3593 CCCAGCTCTCTGGTGCCGAGAAACAGGCAATTTGCCATAGCTGGCTTGTAGACAGC 3652
QY 3603 CTCATATTTTGTGTTTGGATGAGTACATCAGCTGCTGATACAGAAAGTGAAGGTTG 3662
|||||
Db 3653 CTCATATTTTGTGTTTGGATGAGTACATCAGCTGCTGATACAGAAAGTGAAGGTTG 3712
QY 3663 TCCAAAGACCCCTGAGCAAAAGCCAGAGAGGCGCACCTGCATTTGTGATGGCCACGCT 3722
|||||
Db 3713 TCCAAAGACCCCTGAGCAAAAGCCAGAGAGGCGCACCTGCATTTGTGATGGCCACGCT 3772
QY 3723 TGTCCACATCCAGAAATGCAATTAATAGTGTGTTTGAATGTGCAAAAGTCAAGAGC 3782
|||||
Db 3773 TGTCCACATCCAGAAATGCAATTAATAGTGTGTTTGAATGTGCAAAAGTCAAGAGC 3832
QY 3783 ATGGACACATCAACAGCTGCTGGCCCAAGAGCACTAATTTTTCATGATGATGCTCC 3842
|||||
Db 3833 ATGGACACATCAACAGCTGCTGGCCCAAGAGCACTAATTTTTCATGATGATGCTCC 3892
QY 3843 AAGCTGAGCAAAAGCGTAGTAAGCTGTGSCCATATGAGCTGTAAATATTTTAAATAT 3902
|||||
Db 3893 AAGCTGAGCAAAAGCGTAGTAAGCTGTGSCCATATGAGCTGTAAATATTTTAAATAT 3952
QY 3903 TTTGTGTTAAACATGCGATTTAATCAAAAGTTAAAGGTAGAGCACTTACTGGAAAAACTAT 3962
|||||
Db 3953 TTTGTGTTAAACATGCGATTTAATCAAAAGTTAAAGGTAGAGCACTTACTGGAAAAACTAT 4012
QY 3963 GTAGAACTACCTGTTTAACATTTCTGTGCGCAAGTGAAGATCATTTCCCAAGTTCAAG 4022
|||||
Db 4013 GTAGAACTACCTGTTTAACATTTCTGTGCGCAAGTGAAGATCATTTCCCAAGTTCAAG 4072
QY 4023 TCTTCAGATTTTATTAATTAAGGAACCAAAAGCAATTAATCATGATGAATAAATATTTG 4082
|||||
Db 4073 TCTTCAGATTTTATTAATTAAGGAACCAAAAGCAATTAATCATGATGAATAAATATTTG 4132
QY 4083 GNGTAAATGCAATTAATAAATATATAGAGTAATTCAAAGTAGATTTTGTTAATAATTTGA 4142

	gene	/tissue_type="liver"	
		/note="Ivermectin sensitive collies"	
	CDS	7..282	
		/gene="p-gp"	
		7..282	
		/gene="p-gp"	
		/note="abortive product"	
		/codon_start=1	
		/evidence=experimental	
		/product="p-glycoprotein"	
		/protein_id="CAD12806.1"	
		/db_xref="GI:17385399"	
		/translation="MDPEGRGSAEKRNFKMGKSKKNKKPKYSTFAMFRYSN MIDRLVMTGTMAIIHGALPLIMLVFGMTALQMEPDETRFLP"	
	BASE COUNT	1144 a 815 c 964 g 1011 t	
	ORIGIN		
	Query Match	89.4%; Score 3824; DB 4; Length 3934;	
	Best Local Similarity	99.4%; Pred. No. 0;	
	Matches 3850; Conservative	0; Mismatches 20; Indels 4; Gaps 1;	
Db	1	GTGGCGATGCATCCTGTGAAGAGGCCGTAAAGGAGTGCAGAGAACAATTCTTGGAATAAG 70	
	1	GTCACGATGATCTCTGAAGSAGCCGTAGGGGAGTGCAGAGAACAATTCTTGGAAATG 60	
Oy	71	GCGAAAAAAGTAATAAATAATGAGAAGAAAGAAAACCACACTGTCAGACCTTTGCA 130	
Db	61	GCGAAAAAAGTAATAAATAATGAGAAGAAAGAAAACAACACTGTCAGACCTTTGCA 120	
Oy	131	ATGTTTGCCTTTCAAATTTGGCTTGTATAGTTGTATATGTTGTTGGGACAATGGCTGCC 130	
Db	121	ATGTTTGCCTTTCAAATTTGGCTTGTATAGTTGTATATGTTGTTGGGACAATGGCTGCC 180	
Oy	191	ATCATTCATGAGAGTGCACATCCCTCTCCATGANCCTGGTTTTGAAACATGACAGATAGC 250	
Db	181	ATCATTCATGAGAGTGCACATCCCTCTCTCATGATGCTGTTTTGGAAACATGAC----AGC 236	
Oy	251	TTTTGGAATPGCAGGAATTTCAAGAAACAAAACTTTTCAGTTAATTAATGAANGATT 310	
Db	237	TTTTGGAATPGCAGGAATTTCAAGAAACAAAACTTTTCAGTTAATTAATGAANGATT 296	
Oy	311	ACGAACATATACACAACATTTTATCATCAACCATCTGGAGAGAAATGCCACGATGCTCAT 370	
Db	297	ACGAACATATATCACAAATTTTATCATCAACCATCTGGAGAGAAATGACACGATGCTCAT 356	
Oy	371	TATTACATGGGATGGTGTCTGCGCGTGGTGGCTTCATCCAGGTTTCATTTG 430	
Db	357	TATTACATGGGATGGTGTCTGCGCGTGGTGGCTTCATCCAGGTTTCATTTG 416	
Oy	431	TGCTTGGCAGCAGGAAGACAGTACTCAAAATTGAAACAATTTTTTATCTATCATATG 490	
Db	417	TGCTTGGCAGCAGGAAGACAGTACTCAAAATTGAAACAATTTTTTATCTATCATATG 476	
Oy	491	CGACAGAGATTTGGCTGGTTTACCTGCATGACGTTGGGAGCTTAAACACCCGGCTACA 550	
Db	477	CGACAGAGATTTGGCTGGTTTACCTGCATGACGTTGGGAGCTTAAACACCCGGCTACA 536	
Oy	551	GACATGTCTCCAAATTCATATGAAGAAATTTGGCACAAAATTTGGAATGCTTTCACCTCA 610	
Db	537	GACATGTCTCCCAAAATTCATATGAAGAAATTTGGCACAAAATTTGGAATGCTTTCACCTCA 596	
Oy	611	AATACACATTTTTCACCGGTTTTATAGTGGGGTTTACACGTGGTTGGAAAGTAACCTT 670	
Db	597	AATACACATTTTTCACCGGTTTTATAGTGGGGTTTACACGTGGTTGGAAAGTAACCTT 656	
Oy	671	GTCATTTTGGCATTCAGCCCTGTTCTTGACCTTATAGCCCGCATCTGGGAAAGATACTA 730	
Db	657	GTCATTTTGGCATTCAGCCCTGTTCTTGACCTTATAGCCCGCATCTGGGAAAGATACTA 716	
Oy	731	TCTTCATTTACTGTAAGAACCTCTGGGCTATGCAAAAGCTGAGACAGTAGCTGAAGAA 790	
Db	717	TCTTCATTTACTGTAAGAACCTCTGGGCTATGCAAAAGCTGAGACAGTAGCTGAAGAA 776	


```
QY 791 GTCTTAGAGCAATACAGACTGTGATTCCTTTGGAGAGCAAAAGAAAGACTTGAAAG 850
Db 777 GTCTTAGAGCAATACAGACTGTGATTCCTTTGGAGAGCAAAAGAAAGACTTGAAAG 836
QY 891 TACAGCAAAATTTAGAAAGACTTAAAGAAATGGGATTAAGAAAGATTCAGCGCCAAC 910
Db 837 TACAGCAAAATTTAGAAAGACTTAAAGAAATGGGATTAAGAAAGATTCAGCGCCAAC 896
QY 911 ATTTCTATTTGGTCCGCTTCTATGTATGATCTATGATCTATGATCTGCTTCTGAT 970
Db 897 ATTTCTATTTGGTCCGCTTCTTATGATCTATGATCTATGATCTGCTTCTGAT 956
QY 971 GGGACCTCCTTGCTCTCTCCAGTGAATATCTATTTGGACAAAGTACTACTCTCTTT 1030
Db 957 GGGACCTCCTTGCTCTCTCCAGTGAATATCTATTTGGACAAAGTACTACTCTCTTT 1016
QY 1031 TCTGTATTTAATGGGCTTTAGTATGTGACAGGATCCCAAGCATTTGAAGATTGCA 1090
Db 1017 TCTGTATTTAATGGGCTTTAGTATGTGACAGGATCCCAAGCATTTGAAGATTGCA 1076
QY 1091 AACGCAAGAGGAGCAGCTTATGAATCTTCAAGATATTTGCAATTAACCAAGCATTGAC 1150
Db 1077 AACGCAAGAGGAGCAGCTTATGAATCTTCAAGATATTTGCAATTAACCAAGCATTGAC 1136
QY 1151 AGCTATTTGCAAGAGTSGACATTAACCAAGATATTTAAGGAAAAATTGGAATTCAAAAAT 1210
Db 1137 AGCTATTTGCAAGAGTSGACATTAACCAAGATATTTAAGGAAAAATTGGAATTCAAAAAT 1196
QY 1211 GTTCACTTCACTTACCTCTCTCGAAAAAGATTAAGATCTTAAAGGCTTCACCTTAAG 1270
Db 1197 GTTCACTTCACTTACCTCTCTCGAAAAAGATTAAGATCTTAAAGGCTTCACCTTAAG 1256
QY 1271 GTTCAAGAGTGGGAGACAGTGGCGCTGTTGGAAACAGTGGCTCGGGAGAGACAGACC 1330
Db 1257 GTTCAAGAGTGGGAGACAGTGGCGCTGTTGGAAACAGTGGCTCGGGAGAGACAGACC 1316
QY 1331 GTGACAGTGAAGAGGCTCTATGACCCCAAGATGGCATGGTCTATTTGATGACAG 1390
Db 1317 GTGACAGTGAAGAGGCTCTATGACCCCAAGATGGCATGGTCTATTTGATGACAG 1376
QY 1391 GACATTAAGACCATTAATGTAAAGCATCTTCGGGAAATTAAGTGTGATGATCAGAG 1450
Db 1377 GACATTAAGACCATTAATGTAAAGCATCTTCGGGAAATTAAGTGTGATGATCAGAG 1436
QY 1451 CCTGTGTTGTTGGCACCAAGATGCTGAAAAAATTCGCTATGGCCGCAAAATGTCAAC 1510
Db 1437 CCTGTGTTGTTGGCACCAAGATGCTGAAAAAATTCGCTATGGCCGCAAAATGTCAAC 1496
QY 1511 ATGATGATGATTAAGAAAGCTGTTAAGGAAGCAATGCTATGATTTTATCATGAAACTA 1570
Db 1497 ATGATGATGATTAAGAAAGCTGTTAAGGAAGCAATGCTATGATTTTATCATGAAACTA 1556
QY 1571 CCTAATTAATTTGACACTCTGCTGTTGGAGAGAGAGGGCCAGCTGAGTGGTGGACAGAA 1630
Db 1557 CCTAATTAATTTGACACTCTGCTGTTGGAGAGAGAGGGCCAGCTGAGTGGTGGACAGAA 1616
QY 1631 CAGAGAAATCCCAATGCTGCTGGGCCCTGCTGTTGGAAACCCCAAGATTTCTGCTGATGAG 1690
Db 1617 CAGAGAAATCCCAATGCTGCTGGGCCCTGCTGTTGGAAACCCCAAGATTTCTGCTGATGAG 1676
QY 1691 GCMAAGTACACTCTGACACTGAAAGTGAAGAGTGGTTCAAGTGGCCCTGGATTAAGGCC 1750
Db 1677 GCMAAGTACACTCTGACACTGAAAGTGAAGAGTGGTTCAAGTGGCCCTGGATTAAGGCC 1736
QY 1751 AGAAAAGCGGAGCTACCAATTTGATAGTCAATCGTTTGTCTACAGTTCGTAATGCCAT 1810
Db 1737 AGAAAAGCGGAGCTACCAATTTGATAGTCAATCGTTTGTCTACAGTTCGTAATGCCAT 1796
QY 1811 GTGATTTGCTGTTTTGATGATGAGTCAATTTGAGAAAGAAATCATGAGAACTCATG 1870
Db 1797 GTGATTTGCTGTTTTGATGATGAGTCAATTTGAGAAAGAAATCATGAGAACTCATG 1856
QY 1871 AAGAGAAAGGCAATTTACTTCAACTGTGCACATGACAGACAGAGAAATGAATTTGAG 1930
Db 1857 AAGAGAAAGGCAATTTACTTCAACTGTGCACATGACAGACAGAGAAATGAATTTGAG 1916
QY 1931 TTGAAAAATGCCAGTGTGATTCGAAAAAGTGAAGTATGCTTGGAAATGTCTCCAAA 1990
Db 1917 TTGAAAAATGCCAGTGTGATTCGAAAAAGTGAAGTATGCTTGGAAATGTCTCCAAA 1976
QY 1991 GATTCAAGGGTCCAGTTTAATTAATAAGATCAACTCCGAGGATATGACACCCACAA 2050
Db 1977 GATTCAAGGGTCCAGTTTAATTAATAAGATCAACTCCGAGGATATGACACCCACAA 2036
QY 2051 GCGCAAGACAGAAAGCTTGTGTAACAAAAGAGACTGAATGAGAAATGACTCTCAGTTTCC 2110
Db 2037 GCGCAAGACAGAAAGCTTGTGTAACAAAAGAGACTGAATGAGAAATGACTCTCAGTTTCC 2096
QY 2111 TTCTGAGGATTTCTGAAGCTGAACCTCACTGATGAGGCCCTTATTTTGTGTTGTAATTT 2170
Db 2097 TTCTGAGGATTTCTGAAGCTGAACCTCACTGATGAGGCCCTTATTTTGTGTTGTAATTT 2156
QY 2171 TGTGCTATTTAATACGGAGGCCCTGCAACAGCATTTTCAATTAATTTTCAAGATTAATA 2230
Db 2157 TGTGCTATTTAATACGGAGGCCCTGCAACAGCATTTTCAATTAATTTTCAAGATTAATA 2216
QY 2231 GGGATCTTTACCCGAGATGAGGATCTGAAACAAACAGACAGAAATAGTAACATGTTTCT 2290
Db 2217 GGGATCTTTACCCGAGATGAGGATCTGAAACAAACAGACAGAAATAGTAACATGTTTCT 2276
QY 2291 GTATTTGTTTCAAGTCTTGAATTTATTTCTTTATTTTCAATTTTCCGCGGGCTTCACA 2350
Db 2277 GTATTTGTTTCAAGTCTTGAATTTATTTCTTTTATTTTCAATTTTCCGCGGGCTTCACA 2336
QY 2351 TTTGGCAAAAGCTGGGAGAGATCTCTCACTAAGCGGCTTGATACATGTTTTCAGATCCAG 2410
Db 2337 TTTGGCAAAAGCTGGGAGAGATCTCTCACTAAGCGGCTTGATACATGTTTTCAGATCCAG 2396
QY 2411 CTGACACAGATGTCAGCTGTTGATGACCCCTTAACAAACCACTGAGCATTTGACACACC 2470
Db 2397 CTGACACAGATGTCAGCTGTTGATGACCCCTTAACAAACCACTGAGCATTTGACACACC 2456
QY 2471 AGGCTTGCCAAATGATGCGGCTCAAGTTAAAGGGGCTATAGTGTCCAGGCTTGCTGCATT 2530
Db 2457 AGGCTTGCCAAATGATGCGGCTCAAGTTAAAGGGGCTATAGTGTCCAGGCTTGCTGCATT 2516
QY 2531 ACCGAAATATAGCAATCTTGGAGCGGATATATATATCTTAATCTATGTTGGCAA 2590
Db 2517 ACCGAAATATAGCAATCTTGGAGCGGATATATATATCTTAATCTATGTTGGCAA 2576
QY 2591 TTAACACTTTTACTCTTGAACAATTTGACCAATTTGACCAATTTGACAGAGTGTGGAATG 2650
Db 2577 TTAACACTTTTACTCTTGAACAATTTGACCAATTTGACCAATTTGACAGAGTGTGGAATG 2636
QY 2651 AAAATGTTGCTGGACAGACAGCTGAAAGATTAAGAAAGACTTGAAGAGTGGGAAATTT 2710
Db 2637 AAAATGTTGCTGGACAGACAGCTGAAAGATTAAGAAAGACTTGAAGAGTGGGAAATTT 2696
QY 2711 GCTACAGAACCCATCGAAAAAATCTCGAAGCTGTTTCTTATACCGGGAGAGAAAGTTT 2770
Db 2697 GCTACAGAACCCATCGAAAAAATCTCGAAGCTGTTTCTTATACCGGGAGAGAAAGTTT 2756
QY 2771 GAATTAATGATGATCAGACAGAGTTTGAAGTACATACAGAAACTCTTTGAGGAAAGCACAC 2830
Db 2757 GAATTAATGATGATCAGACAGAGTTTGAAGTACATACAGAAACTCTTTGAGGAAAGCACAC 2816
QY 2831 ATCTTGCGGGGTCTCATTTCTATCACCCCGGCAATGATATTTTCTCTATGCTGGCTGT 2890
Db 2817 ATCTTGCGGGGTCTCATTTCTATCACCCCGGCAATGATATTTTCTCTATGCTGGCTGT 2876
QY 2891 TTCCGGTTTGTGCTCTACTTGTGSCAAATGATGATCAATCTTTCAGAGATGTTCTTTTG 2950
Db 2877 TTCCGGTTTGTGCTCTACTTGTGSCAAATGATGATCAATCTTTCAGAGATGTTCTTTTG 2936
QY 2951 GTATTTCTCAGTATTTGTCTTGTGTCATGCGAGTGGGCGAGGTCAATTTGCTCTCT 3010
Db 2937 GTATTTCTCAGTATTTGTCTTGTGTCATGCGAGTGGGCGAGGTCAATTTGCTCTCTCT 3010
```

Db 2937 GTATTCTACGATTTCTTGTGGCCATGCGAGTGGGGAGGTAGTTCATTGCTCCT 2996
Oy 3011 GACTATGCCAAAGCCAAAGATATGACGAGCCCGATCATGATCTTTAAAAAGCCCT 3070
Db 2997 GACTATGCCAAAGCCAAAGATATGACGAGCCCGATCATGATCTTTAAAAAGCCCT 3056
Oy 3071 CTGATTGACAGCTACAGCCCTCAGCGGCTCAGCAAAATTCATTTGAGAGAAATGTGACA 3130
Db 3057 CTGATTGACAGCTACAGCCCTCAGCGGCTCAGCAAAATTCATTTGAGAGAAATGTGACA 3116
Oy 3131 TTTAATGAGTCTGTTTCAACTATCCACTGACAGACATCCCGTCTCCAGGGCTG 3190
Db 3117 TTTAATGAGTCTGTTTCAACTATCCACTGACAGACATCCCGTCTCCAGGGCTG 3176
Oy 3191 AGCCTGAGGTGAGAGAGGGCCCAAGCTGGCCCTCTAGTATGACAGTGGCTGGGAAG 3250
Db 3177 AGCCTGAGGTGAGAGAGGGCCCAAGCTGGCCCTCTAGTATGACAGTGGCTGGGAAG 3236
Oy 3251 AGCAGAGTGTTCAGCTCTAGAGCGCTTCTATGACCCCTTGGCTGGTTCAGTCTAAT 3310
Db 3237 AGCAGAGTGTTCAGCTCTAGAGCGCTTCTATGACCCCTTGGCTGGTTCAGTCTAAT 3296
Oy 3311 GATGGCAAGAGATTAAGACCTGATATGTCAGTGGCTCCGAGACACCTGGGATGCTG 3370
Db 3297 GATGGCAAGAGATTAAGACCTGATATGTCAGTGGCTCCGAGACACCTGGGATGCTG 3356
Oy 3371 TCTCAGAGACCCATCTGTTTGTGACAGCATTTGCCGAGAACATTTGCTATGAGACAC 3430
Db 3357 TCTCAGAGACCCATCTGTTTGTGACAGCATTTGCCGAGAACATTTGCTATGAGACAC 3416
Oy 3431 AGCCGGGCTGTATCATATGAGATTTATGACAGAGCAGAGAGAGGACCATCATCACAC 3490
Db 3417 AGCCGGGCTGTATCATATGAGATTTATGACAGAGCAGAGAGAGGACCATCATCACAC 3476
Oy 3491 TTGATGAGAGACTCCCTGAGAAATACAAACACAGAGTAGAGAGCAAGAAAGAACCCAGCTC 3550
Db 3477 TTGATGAGAGACTCCCTGAGAAATACAAACACAGAGTAGAGAGCAAGAAAGAACCCAGCTC 3536
Oy 3551 TCTGTGGCCAGAAAGAGCGCATTTGCCATAGCTCCGCTCTTGTATGACAGCTCATATT 3610
Db 3537 TCTGTGGCCAGAAAGAGCGCATTTGCCATAGCTCCGCTCTTGTATGACAGCTCATATT 3596
Oy 3611 TTGCTTTGGATTAAGACTCATATGATGATGATGATGATGATGATGATGATGATGATGAT 3670
Db 3597 TTGCTTTGGATTAAGACTCATATGATGATGATGATGATGATGATGATGATGATGATGAT 3656
Oy 3671 GCCCTGACAAAGCCAGAGAGAGCCGACCTGATTTGATGATGATGATGATGATGATGATGAT 3730
Db 3657 GCCCTGACAAAGCCAGAGAGAGCCGACCTGATTTGATGATGATGATGATGATGATGATGAT 3716
Oy 3731 ATCCAGATGCAATTTAATAGTGTGTTTCAATGCGAAAGTCAAGAGCATGGACACA 3790
Db 3717 ATCCAGATGCAATTTAATAGTGTGTTTCAATGCGAAAGTCAAGAGCATGGACACA 3776
Oy 3791 CATCAACAGCTGCTGCCCAAGAAAGCATTTTTCATGCTGATGCTGATGCTGATGCTGAT 3850
Db 3777 CATCAACAGCTGCTGCCCAAGAAAGCATTTTTCATGATGATGATGATGATGATGATGAT 3836
Oy 3851 GCAAAAGCGTAGTAAGTGTGGCATATGAGCTG 3884
Db 3837 GCAAAAGCGTAGTAAGTGTGGCGCGCGCTG 3870

RESULT 9
LOCUS 108557 108557 4669 bp DNA linear PAT 02-DEC-1994
DEFINITION Sequence 3 from Patent WO 87/05943.
ACCESSION 108557
VERSION 108557.1 GI:588735
KEYWORDS
SOURCE Unknown.
ORGANISM Unknown.
Unclassified.

REFERENCE 1 (bases 1 to 4669)
AUTHORS Roninson,I.B., Pastan,I.H. and Gottesman,M.M.
TITLE COMPOSITIONS AND METHODS FOR CLONES CONTAINING DNA SEQUENCES
JOURNAL ASSOCIATED WITH MULTIDRUG RESISTANCE IN HUMAN CELLS
FEATURES Patent: WO 8705943-A 3 08-OCT-1987;
Source Location/Qualifiers
1..4669
BASE COUNT 1394 a 892 c 1129 g 1254 t
ORIGIN
Query Match 78.1%; Score 3342.8; DB 6; Length 4669;
Best Local Similarity 87.6%; Pred. No. 0;
Matches 3759; Conservative 0; Mismatches 487; Indels 44; Gaps 8;
Oy 1 GGAGCGGAGGTCCGGATGATGATCTGTAAGAGAGCCCTTAAGGGAGAGAGAGAACTT 60
Db 409 GGAGCGGAGGTCCGGATGATGATCTGTAAGAGAGAGCCCTTAAGGGAGAGAGAGAA 468
Oy 61 CTGGAATAATGGGCAAAAAAGTAAAAAATGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 120
Db 469 CTTTTTAACTGAACATTAATAAGTGAAGAAATGAGAGAGAGAGAGAGAGAGAGAGAG 528
Oy 121 CACGTTTCAATGTTTGGCTTTCATTAATGCTGATAGGTTGATATGTTGGGAGAC 180
Db 529 TGTATTTTCAATGTTTGGCTTTCATTAATGCTGATAGGTTGATATGTTGGGAGAG 588
Oy 181 AATGGCTGCATCATCATGAGAGCTGCACCTCCCTCATGATGCTGTTTGGAAACAT 240
Db 589 TTTGGCTGCATCATCATGAGAGCTGCACCTCCCTCATGATGCTGTTTGGAAACAT 648
Oy 241 GACAGATACCTTTGCAATTCAGAGAAATTCAGAGAGAGAGAGAGAGAGAGAGAGAGAG 300
Db 649 GACAGATACCTTTGCAATTCAGAGAAATTCAGAGAGAGAGAGAGAGAGAGAGAGAGAG 702
Oy 301 TGAAGTATTTAGCAACAATTCACATTTTCATGACCATCTGAGAGAGAGAGAGAGAGAT 360
Db 703 TGAAGTATTTAGCAACAATTCACATTTTCATGACCATCTGAGAGAGAGAGAGAGAGAT 762
Oy 361 GTATGGCTATTTTACAGTGGAGTGGTGGCTGGCTGGCTGGCTGGCTGGCTGGCTGGCTGG 420
Db 763 GTATGGCTATTTTACAGTGGAGTGGTGGCTGGCTGGCTGGCTGGCTGGCTGGCTGGCTGG 822
Oy 421 TTGATTTCTGGTGGCTGGAG 480
Db 823 TTGATTTCTGGTGGCTGGAG 882
Oy 481 TGTATCATGCGACAGAGATGCTGTTGATGATGATGATGATGATGATGATGATGATGATGAT 540
Db 883 TGTATCATGCGACAGAGATGCTGTTGATGATGATGATGATGATGATGATGATGATGATGAT 942
Oy 541 CCGGCTCAGAGAGATGCTCCAAATCAATGAAGAAATGGCGCAAAATTTGAAATGTT 600
Db 943 CCGACTTACAGAGATGCTCTTAAGATTAATGAAGTATTTGTTGACAAATTTGAAATGTT 1002
Oy 601 CTTTCACTCAATGAACAATTTTTCACCGTATTTATAGTGGGTTTACAGTGGTTGAA 660
Db 1003 CTTTCACTCAATGAACAATTTTTCACCGTATTTATAGTGGGTTTACAGTGGTTGAA 1062
Oy 661 GCTAACCTTGTGATTTTGGCATCAGCCGTGTTCTTGAAGCTTTCAGCCGATGGGC 720
Db 1063 GCTAACCTTGTGATTTTGGCATCAGCTCTGTTTGAAGCTTTCAGCTGCTGCTGGCC 1122
Oy 721 AAAGATCATCTTCAATTTACTGATTAAGAACTTGGCTATGCAAAAGCTGAGCACT 780
Db 1123 AAAGATCATCTTCAATTTACTGATTAAGAACTTGAAGCACTTGAAGCAAAAGCTGAGCACT 1182
Oy 781 AGCTGAAGAGTCTTACAGCAATCAGAACTGTATGCTTGTGAGAGCAAAAGAGAA 840
Db 1183 AGCTGAAGAGTCTTGTGAGCAATTTAGAACTGTATGCTTGTGAGAGCAAAAGAGAA 1242
Oy 841 ACTGAAAGGTACAACAAAATTTAGAAAGAGCTAAAGAAATGGAGTAAGAAAGACTAT 900

Db 1243 ACTGAAAGGTACAACAAAATTTAGAGAAGCTAAAGAAATTGGATAAAGAAAGCTAT 1302
Qy 901 CACGGCCAAACATTTCTATGTGGCCGCTTCTTATGTATGTATGATATGATATGCTCTGGC 960
Db 1303 TACAGCCAAATTTCTATGTAGGTGCTGCTTCCCTGCTGATCTATGATCTTATGCTCTGGC 1362
Qy 961 TTTTGTGTATGGACCTCTTGGTCTCTCCAGTGAATATTTATTTGGACAGTACTGAC 1020
Db 1363 CTTCGTGTATGGACCACTTGGTCTCTCAGGGGAATTTCTATTTGGACAGTACTGAC 1422
Qy 1021 TGTCTTTCTTTCTGTATTTATTTGGGCTTTTATGTATTTGACAGGATCCCAAGCATTTGA 1080
Db 1423 TGTATTTCTTTCTGTATTTATTTGGGCTTTTATGTATTTGACAGGATCTCCAGCATTTGA 1482
Qy 1081 AGCATTTGCAAAACGCAAGAGAGAGACCTTATGAAATCTTCAAGATTAATTTGACATTAACC 1140
Db 1483 AGCATTTGCAAAATGCAAGAGAGAGACCTTATGAAATCTTCAAGATTAATTTGAAATTAACCC 1542
Qy 1141 AAGCATTTGACAGTATTTCCAAAGATGTGACATTAACCCAGATTAATTTAAGGAAATTTGGA 1200
Db 1543 AAGTATTTGACAGTATTTCCAAAGATGTGACACAAACCAAGATTAATTTAAGGAAATTTGGA 1602
Qy 1201 ATTCAAAATTTGTCACCTTATGATTTACCTTTCTGAAAGAAAGATTAAGATCTTAAAGGCTCT 1260
Db 1603 ATTCAGAAATTTGTCACCTTATGATTTACCTTTCTGAAAGAAAGATTAAGATCTTAAAGGCTCT 1662
Qy 1261 CAACCTGAAGGTTTCAGAGTGGGCGACAGAGTGGCGCTGGTTGGGAACAGTGGCTCGGGAA 1320
Db 1663 GAACCTGAAGGTTTCAGAGTGGGCGACAGAGTGGCGCTGGTTGGGAACAGTGGCTCGGGAA 1722
Qy 1321 GAGCAGCAGCCGTGACGTGTATGAGAGGCTGTATGACCCCAACAGATGGCATGGTCTGTAT 1380
Db 1723 GAGCAGCAACAGTCCAGCTGTATGAGAGGCTGTATGACCCCAACAGAGGAGTGGTCTGTAT 1782
Qy 1381 TGATGCAAGAGCATTTAGGACCATTAATGTAAAGCATCTTGGGAAATTTATGCTGTGTGT 1440
Db 1783 TGATGCAAGAGCATTTAGGACCATTAATGTAAAGCATCTTGGGAAATTTATGCTGTGTGT 1842
Qy 1441 GAGTGAAGAGCCTGTGTTTGGCCACAGATAGCTGGAAGCAATTCGCTGATGGCCGCA 1500
Db 1843 GAGTGAAGAGCCTGTGTTTGGCCACAGATAGCTGGAAGCAATTCGCTGATGGCCGCA 1902
Qy 1501 AAATGTCCATGTGATGATTTGAAAGCTGTAAAGGCAAGCAATTCGCTGATGATTTAT 1560
Db 1903 AAATGTCCATGTGATGATTTGAAAGCTGTCAAGGAAAGCAATTCGCTGATGATTTAT 1962
Qy 1561 CATGAAATCTACCTAATTAATTTGACACTCTGTTGAGAGAGAGGGGCCAGCTGATGG 1620
Db 1963 CATGAAATCTACCTAATTAATTTGACACTCTGTTGAGAGAGAGGGGCCAGCTGATGG 2022
Qy 1621 TGGACAGAAACAGAGATGCCATTTGCTGGGGCCGCTGTTGCAACCCCAAGATCTCTCT 1680
Db 2023 TGGACAGAAACAGAGATGCCATTTGCTGGGGCCGCTGTTGCAACCCCAAGATCTCTCT 2082
Qy 1681 GCTGAGTGAAGGCAAGTCCAGCTGTGACACTGAAGTGAAGAGTGTTCAGAGTGGCCCT 1740
Db 2083 GCTGAGTGAAGGCAAGTCCAGCTGTGACACTGAAGTGAAGAGTGTTCAGAGTGGCCCT 2142
Qy 1741 GGATTAAGGCCAGAAAGGCCGACTCCATTTGTGATAGCTACTGCTTTGTCTACAGTTGC 1800
Db 2143 GGATTAAGGCCAGAAAGGCCGACTCCATTTGTGATAGCTACTGCTTTGTCTACAGTTGC 2202
Qy 1801 TAATGCCGATGTCTGCTGCTTTGTGATGATGAGTCAATTTGAGAGAAAGAAATTCATGA 1860
Db 2203 TAATGCCGATGTCTGCTGCTTTGTGATGATGAGTCAATTTGAGAGAAAGAAATTCATGA 2262
Qy 1861 TGAATCTAGAAAGAGAGAGGATTTACTTCAAACTTGTCCAAATGCAAGACAGAGAAAA 1920
Db 2263 TGAATCTAGAAAGAGAGAGGATTTACTTCAAACTTGTCCAAATGCAAGACAGAGAAAA 2322
Qy 1921 TGAATTTAGTTAGAAAAATGCCACTGTGTGAATCCAAAAAGTGAAGTATGCTGTGAAT 1980
Db 2323 TGAATTTAGTTAGAAAAATGCCACTGTGTGAATCCAAAAAGTGAAGTATGCTGTGAAT 2382

Qy 1981 GTCTCAAAAGATTCAGAGGTCACGTTTAATTAAGAAAGATCAACTGCGAGAGATATCA 2040
Db 2383 GTCTTCAAAAGATTCAGAGGTCACGTTTAATTAAGAAAGATCAACTGCGAGAGATATCA 2442
Qy 2041 TGCACCAAGAGGCCCAAGACAGAAAGCTGTGTACAAAAGAGGACTGAATGAAATGTACC 2100
Db 2443 TGCATCACAAGGCCCAAGACAGAAAGCTGTGTACAAAAGAGGCTGTGTGAATGAATGTACC 2502
Qy 2101 TCCAGTTTCTTTCTGAGAGATTCGAGAGTCAACTGAATGAGCCCTTAATTTGTGTGT 2160
Db 2503 TCCAGTTTCTTTCTGAGAGATTCGAGAGTCAACTGAATGAGCCCTTAATTTGTGTGT 2562
Qy 2161 TGTATATTTTGTCTATTTAATTAAGAGGCTGTCAACCCAGATTTTAATTAATTTTC 2220
Db 2563 TGTGTATTTTGTCTATTTAATTAAGAGGCTGTCAACCCAGATTTTAATTAATTTTC 2622
Qy 2221 AAGGATTAAGGATCTTTTACCAGATGAGATTCCTGAACCAAGACAGATTAATTAAT 2280
Db 2623 AAGGATTAAGGATCTTTTACCAGATGAGATTCCTGAACCAAGACAGATTAATTAAT 2682
Qy 2281 CATGTTTCTGATTTGTTTCTAGTCCCTGGAATTTCTTTTATACATTTTCTCTCA 2340
Db 2683 CTGTTTCTGATTTGTTTCTAGTCCCTGGAATTTCTTTTATACATTTTCTCTCA 2742
Qy 2341 GGGCTTCACTTTTGGCAAAAGCTGGGAGATCTCTCAAGGCGCTTGATATACATGTTT 2400
Db 2743 GGGCTTCACTTTTGGCAAAAGCTGGGAGATCTCTCAAGGCGCTTGATATACATGTTT 2802
Qy 2401 CAGATCCATGCTGAGACAGATGTCAGTGGTTGATGACCTTAAGAAACCAACTGGAG 2460
Db 2803 CAGATCCATGCTGAGACAGATGTCAGTGGTTGATGACCTTAAGAAACCAACTGGAG 2862
Qy 2461 ATTGCAACCAAGGCTTGCAATGATGCGGCTCAAGATTAAGGGGCTATAGSTTCCAGGCT 2520
Db 2863 ATTGCAACCAAGGCTTGCAATGATGCGGCTCAAGATTAAGGGGCTATAGSTTCCAGGCT 2922
Qy 2521 TGCCTATTTCCCAAGAAATATAGCAAACTTGGAGACAGCATTAATTAATCCCTTAATCTA 2580
Db 2923 TGCCTATTTCCCAAGAAATATAGCAAACTTGGAGACAGCATTAATTAATCCCTTAATCTA 2982
Qy 2581 TGTGTGCAATTAACATTTTACTTTAGCAATTTGACCATTCATTCATTCATTCATTCAT 2640
Db 2983 TGTGTGCAATTAACATTTTACTTTAGCAATTTGACCATTCATTCATTCATTCATTCAT 3042
Qy 2641 TGTGAAATTAATAATGTTGCTGAGACAGCACTGAAGATTAAGAAAGAGCTGAAGAGAGC 2700
Db 3043 TGTGAAATTAATAATGTTGCTGAGACAGCACTGAAGATTAAGAAAGAGCTGAAGAGAGC 3102
Qy 2701 TGGGAAGATTTGTACAGAAAGCCATTCGAAAACTTCCGAACCTGTTTCTTTGATCGGCA 2760
Db 3103 TGGGAAGATTTGTACAGAAAGCCATTCGAAAACTTCCGAACCTGTTTCTTTGATCGGCA 3162
Qy 2761 GCGAAGTTTGAATATCATGTATGACAGAGTGTGCAAGTACCATTCAGAAACTCTTTGAG 2820
Db 3163 GCGAAGTTTGAATATCATGTATGACAGAGTGTGCAAGTACCATTCAGAAACTCTTTGAG 3222
Qy 2821 GAAGACACATCTTGGGGGCTCAATTTTCTATTCACCCAGCATGATGTATTTTCTTA 2880
Db 3223 GAAGACACATCTTGGGGGCTCAATTTTCTATTCACCCAGCATGATGTATTTTCTTA 3282
Qy 2881 TGTGTGCTTTTCCGTTTGTGTGCTACTTGTGTGCAAAATGATGATGATGATGATGAT 2940
Db 3283 TGTGTGCTTTTCCGTTTGTGTGCTACTTGTGTGCAAAATGATGATGATGATGATGAT 3342
Qy 2941 TGTGCTTTTGTATTTCTAGCATTTGCTTTGTGTGCTAGTGGCAAGTGGCAAGTGGCA 3000
Db 3343 TGTGCTTTTGTATTTTGTAGCTGTTGTGTGTGCTAGTGGCAAGTGGCAAGTGGCA 3402
Qy 3001 ATTGCTCTGATATGCAAAAGCCAAAGTATCAGAGCCAGCTCATCATGATCATGATGA 3060
Db 3403 ATTGCTCTGATATGCAAAAGCCAAAGTATCAGAGCCAGCTCATCATGATCATGATGA 3462

```

OY 3061 AAAAGCCCTGATGACAGCTACAGCCCTCAGGCTCAAGCAAAATACGTGAAG 3120
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 3463 AAAAACCCTTTGATGACAGCTACAGCAAGCCCTAAAGCCGAAACATTTGAAG 3522
OY 3121 AAATGCAATTTAATGAGCTGTGTCAACTATCCACTGACAGACATCCCGTGT 3180
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 3523 AAATGTCACATTTGATGAGAGTTGTATCAACTATCCACCGAGCGGACATCCAGTGT 3582
OY 3181 CCAGGGCTGAGGCTGAGGTGAAGAGGCGCAGACGCTGGCCCTCGTAGAGTACAGTG 3240
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 3583 TCAGGGAGCTGAGGCTGAGGTGAAGAGGCGCAGACGCTGGCTGTGTGGGAGCAGTG 3642
OY 3241 CTGTGGAGAGACAGAGTTGTTCAGCTCTAGAGCGCTCTATGACCCCTTGGCTGTTC 3300
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 3643 CTGTGGAGAGACAGAGTTGTTCAGCTCTAGAGCGGCTCTAGAGCCCTTGGAGGAA 3702
OY 3301 AGTCTAATTTGATGCAAAAGATTAAGACCTGAATGTCCAGTGTCCGAGCAGACCT 3360
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 3703 AGTCTCTGTTGATGCAAAAGATTAAGAGCAATGTTCAGTGTCCGAGCAGACCT 3762
OY 3361 GGGCATGCTGTCTAGAGCCCATCCTGTTTGAAGCTGACAGATTGCCAGACATTGGCTA 3420
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 3763 GGGCATGCTGTCTAGAGCCCATCCTGTTTGAAGCTGACAGATTGCTGAGACATTGGCTA 3822
OY 3421 TGGAGACAAGAGCGGGTCTATGCATGGAAGATTATGACAGCAGCAGAGGCCAA 3480
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 3823 TGGAGACAAGAGCGGGTGTGTCTACAGGAAGATGTGAGGGCAGCAAGAGAGGCCAA 3882
OY 3481 CATACACCCTTCATGACAGACATCCTCTGAGAAATACACACAGATGAGACAAAG 3540
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 3883 CATACATCCTTCATGACAGACATCCTCTGAGAAATACATGATGAGAGACAAAG 3942
OY 3541 AACCAGCTCTCTGTGGCCAGAAACAGCCGATGCCATGCTCGGCTCTTGTAGACA 3600
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 3943 AACTAGCTCTGTGGCCAGAAACAGCATTTGCCATGCTCGCTGTGTAGACA 4002
OY 3601 GCCTCATATTTTGTCTTTTGTGATGAGCTACATCAGCTCTGATACAGAAAGTGAAGGT 3660
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 4003 GCCTCATATTTTGTCTTTGATGAGAGCCAGCTCTGATACAGAAAGTGAAGGT 4062
OY 3661 TGTCCAGAGAGCCGCGGAGAAAGCCAGAGAGCCGACCTGATGTGATCGCCACCG 3720
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 4063 TGTCCAGAGAGCCGCGGAGAAAGCCAGAGAGCCGACCTGATGTGATGTGCTACCG 4122
OY 3721 CTGTCTCACCATCGAATGAGATTAATAGTGTGTTTACAGATGCGCAAGCAAGCA 3780
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 4123 CTGTCTCACCATCGAATGAGATTAATAGTGTGTTTACAGATGCGCAAGCAAGCA 4182
OY 3781 GCATGCGACACATCGAAGACCTGCTGGCCAGAAAGCATCTATTTTTCATGTGAGTGT 3840
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 4183 GCATGCGACACATCGAAGACCTGCTGGCCAGAAAGCATCTATTTTTCATGTGAGTGT 4242
OY 3841 CCAGGGCTGAGCAAGAGCTAGTAAGCTGTGSCATATGAGCTGTAAATTTTAAAT 3900
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 4243 CCAGGGCTGAGCAAGAGCGCAAGTAAGCTGTGATGATGATGTTAAATCTTTTAAAT 4302
OY 3901 ATTTGTTTAAACATGAGCTTAAATCAAGATTAAGGTGAAGCTTACGCAAGAAACT 3960
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 4303 ATT--TGTTAGATATGACATTTTATTCAGAGTAA--AGCAACACTTACGAAT 4355
OY 3961 ATGTAGAACATGCTTTTACATTTCTGTGCACTGACATGAGATCATCCACCAAGTTCAG 4020
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 4356 ATGTAGAACATGCTTTTACATTTCTGTGCACTGACATGAGATCATCCACCAAGTTCAG 4403
OY 4021 AGTCTCAGATTTTAAATTAAGAAACAA--AAGAAACATTAATGATGATGATTAAT 4078
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 4404 -----AGAGACTTCGTAATTAAGAAACAGAGTGAAGACATCATCAAGTGAAGAAAT 4458
OY 4079 ACTGCTGTATTTGCAATTAATTAATTAAGATTAATTCAGATTAATTCAGATTAATTAAT 4138
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 4459 CATAGTTTAACTCATTAATTAATTAATTAATTAATTAAGATTAATTAAGATTAATTAAT 4518
OY 4139 TGTATTAATTTTGTATTAATTTT-----ATTGTAAGTACTGCTTTCGTAAGAAAT 4190

```

```

Db 4519 ATGTGTAATTTTGTATTAATTTTCCATTTTGACTGTAACTGACCTGCTAAAGAT 4578
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
OY 4191 TATAGAGTGTAAAGTACTG--AAGTTGAATTAAGCTCTGATTAATTAAGTAA 4249
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 4579 TATAGAGTGTAAAGTAAATTAATTTGCAATTAAGTGT--TCTATTAATTAAGTAA 4635
OY 4250 ACTTTTATTCAAAAAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 4279
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 4636 ACTTTCATGTGAAAAAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 4665

RESULT 10
E02326 4378 bp RNA linear PAT 29-SEP-1997
LOCUS Multidrug resistance relating gene derived from human normal cells.
DEFINITION E02326
ACCESSION E02326.1 GI:2170561
VERSION JP 1990100680-A/1.
KEYWORDS Homo sapiens.
SOURCE Homo sapiens.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominiidae; Homo.
REFERENCE
  1 (bases 1 to 4378)
  Ueda,K. and Komano,T.
  HUMAN NORMAL CELL-DERIVED MDR RELATED GENE
  Patent: JP 1990100680-A 1 12-APR-1990.
  JOURNAL SUNTORY LTD

COMMENT
OS Homo sapiens
PN JP 1990100680-A/1
PD 12-APR-1990
PF 05-OCT-1988 JP 1988251475
PI UEDA KAZUMITSU, KOMANO TORU
PC C12N15/12,C12N1/21,C12O1/68;
CC strandedness: Single;
CC topology: Linear;
CC *source: tissue-type=Adrenal gland;
CC *source: clone=SM1132;
FH key Location/Qualifiers
FT 5'UTR 1..137
FT CDS 138..3980
FT mat_peptide 138..3977
FT /gene='Multidrug resistance relating gene' FT

FEATURES
  source FT 3'UTR /gene='MDRA1'
  Location/Qualifiers
  1..4378
  /organism='Homo sapiens'
  /db_xref='taxon:9606'
  1188 t

BASE COUNT 1315 a 818 c 1057 g 1188 t

ORIGIN
Query Match 78.1%; Score 3341.2; DB 6; Length 4378;
Best Local Similarity 87.6%; Pred. No. 0;
Matches 3758; Conservative 0; Mismatches 488; Indels 44; Gaps 8;

```

QY	241	GAAGAGTACCTTGGAAATGACAGAACTTTCAGGAAGACAAACCTTTCAGATTAAATTA	300
Db	362	GACAGATCTCTTGGAAATGACAGAAATTTAAGAAATCTGTGTG-----CAACATCAC	415
QY	301	TGAAGTATTAAGAAACATACACACATTTTCATCAACCATCTGGAGAGAAATGACAC	360
Db	416	TAATTAAGATGATATTCATCATATACAGGGTCTTTCATGTGATCTGGAGAGACATACAG	475
QY	361	GTATGCGCTTATTAACAGTGGATCGSTCTGCGCTGCTGGTGGCTCTTAATCCAGAT	420
Db	476	GTATGCGCTTATTAACAGTGGAAATGTGCTGCTGGTGGCTGCTTCAATTCAGAT	535
QY	421	TTGATCTGGGCGCGGACAGAGAGACAGATACACTCAAAATTTGAAACAAATTTTTC	480
Db	536	TTGATTTTGTGTCCTGGACGCTGGAAACAAATACAAATTTGAAACAAAGTTTTTTC	595
QY	481	TGCTATCATGCGACAGAGATTTGCTGTGGTTGACSTGCATGACGTTGGAGCTTAAC	540
Db	596	TGCTATAAGCGACAGAGATAGCGCTGGTTGATGTGCACATTTGGAGAGCTTAAC	655
QY	541	CCGCGTCACAGACGATGTCTCCAAAATCATATAAGAAATTTGGCAGCAAAATTTGATGT	600
Db	656	CCGACTTACAGATGATGTCTCCAAAGATTAATGAAGAAATTTGTGACAAAATTTGATGT	715
QY	601	CTTGACCTCAAGAACATATTTTCCACGSGTTTATATGGGGTTTACACGTGGTTGAA	660
Db	716	CTTGACGTCAAGAACACATTTTTCACGTGGTTTATATAGATTTTACAGTGGTTGAA	775
QY	661	GCTAACCCCTGTGATTTTGGCCATACACCCTGTCTTGGACTTTCACGCGCATCTGGGC	720
Db	776	GCTAACCCCTGTGATTTTGGCCATACAGTCCCTGTCTTGGACTTTCACGCTGTCTGGGC	835
QY	721	TAAGATACATATTTCTATTTACTGATTAAGAACTCTTGGCTATGCAAAACCTGAGACGT	780
Db	836	TAAGATACATATTTCTATTTACTGATTAAGAACTCTTACGATGCAAAACCTGAGACGT	895
QY	781	AGCTGAAGAGCTTAGAGCAATACGAAGCTGATGTCCTTGGAGAGCAAAAGAAAGA	840
Db	896	AGCTGAAGAGCTTGGCAGCAATAGAACGTGATTTGATTTGGAGAGCAAAAGAAAGA	955
QY	841	ACTTGAAGGTACAACAAAATTTAGAGAGCTTAAGGAATTTGGATTAAGAAAGCTAT	900
Db	956	ACTTGAAGGTACAACAAAATTTAGAGAGCTTAAGGAATTTGGATTAAGAAAGCTAT	1015
QY	901	CACGCGCAACATTTCTATTTGTTGGTCCCTTTCTATTTGATCTATGACATATCTCTGGC	960
Db	1016	TACAGCAATATTTCTATTTAGTGTGCTTCTTCCGTGATCATCATCTTATCTCTGGC	1075
QY	961	TTTCTGATTTGGAGACCTCTGTGCTGCTCCAGTAAATATCTTTGAGAAATGCTAC	1020
Db	1076	CTTCTGATTTGGAGACCACTGTGTCTCTCAGAGGAAATTTCTATTTGAGAAATGCTAC	1135
QY	1021	TGTCCTCTTCTGTATTAATTGGGGCTTTTATGATTTGACAGGCATCCCCAAGCTTGA	1080
Db	1136	TGATTTCTTCTTGTATTAATTGGGGCTTTTATGATTTGACAGGCATCCCCAAGCTTGA	1195
QY	1081	AGCATTTGCAAAACGAGAGAGACAGCTTATATAATCTTCAAGATTAATGACATTAAC	1140
Db	1196	AGCATTTGCAAAATGCAAGAGAGAGCTTATATAATCTTCAAGATTAATGATTAATAGCC	1255
QY	1141	AAGCATTTGACACTTCTGGAAGAGTGGACATTAACACATATTTTAAGGGAATTTGGA	1200
Db	1256	AAGTATTTGACACTTCTGGAAGAGTGGACCAACACATATTTTAAGGGAATTTGGA	1315
QY	1201	ATTCAAAAATGTTCACTTCAGTTACCTTCTGAAAAGAGATTAAGATCTTAAGGGCT	1260
Db	1316	ATTCAAAAATGTTCACTTCAGTTACCATCTCTGAAAAGAGATTAAGATCTTAAGGGCT	1375
QY	1261	CAACCTGAAGGTTGAGAGTGGGAGACAGTGGCTGTGGAGAGAGTGGCTGGCGGGA	1320
Db	1376	GAACCTGAAGGTTGAGAGTGGGAGACAGGCTGGCTGTGGAGAGAGTGGCTGGCGGGA	1435

QY	1321	GGAGACCGACGGTGCAGTGAACGAGAGGGCGTATACCCGACACATGGCATGGTGGTAT	1380
Db	1436	GAGCACAACAGTCCAGCGTATGACGAGGCGCTCTATACCCGACAGAGGGATGGTCAAGTCT	1495
QY	1381	TGATGACACAGGACATTTAGGACCAATAAATGTAAAGCATCTTCGGAAATTAATCTGGTGTG	1440
Db	1496	TGATGACACAGGATATTTAGGACCAATAAATGTAAAGGTTTTCTAGGGAAATCAATTCGTGTGTG	1555
QY	1441	GAGTCAGGAGCCGTGTGTTGGCCACACAGATACCTGAATAACATTCGTATGAGCGCGGA	1500
Db	1556	GAGTCAGGAACTGTATGTTTGGCACACAGATAGCTGAATAACATTCGTATGAGCGCGGA	1615
QY	1501	AAATGTCCACATGATGTGATGTGAAGAAAGGTGTAAAGAACCCAAATCCGTATGATTTAT	1560
Db	1616	AAATGTCCACATGATGTGATGTGAAGAAAGGTGTCAAGAGAACCCAAATCCGTATGATTTAT	1675
QY	1561	CATGAACACTACCTAATTAATTTGACACTGTGTTGAGAGAGAGGGGCCGACTGAGTGG	1620
Db	1676	CATGAACACTCCTCATTAATTTGACACCCGTGTTGAAGAGAGGGGCCGACTGAGTGG	1735
QY	1621	TGCGAGAAACAGAAATTCGCGATTCGTGGGCGCCGTGGTTCGAACCCCAAGATTCCTCT	1680
Db	1736	TGCGCGAAGACAGAGATTCGCGATTCGCGGCGCCGTGGTTCGAACCCCAAGATTCCTCT	1795
QY	1681	GGTGGATGAGGACAGCTGACGTCTGGACACTGAAGTAAGCAAGTGGTTCAGTGGCCCT	1740
Db	1796	GCTGGATGAGGCGACGTGACGCTTGGACACAGAAACGAAAGCATGGTTCAGTGGCCCT	1855
QY	1741	GGATTAAGGCCAGAAAAAGGCCGAGCTACCATTTGATAGCTCATCTGTTTGTCTACAGTTGG	1800
Db	1856	GGATTAAGGCCACAAAAAGGCTCGGACACACATTTGATAGCTCATCTGTTTGTCTACAGTTGG	1915
QY	1801	TAAATGCCGATGTGCATTTGCTGTGTTTGTGATGTGAGAGCATTCATTCGAGAAAGGAATATGA	1860
Db	1916	TAAATGCTGACGTTCATTCGCTGGTTTTGATGTGATGAGATCATTCGTGGAGAAAGGAATATGA	1975
QY	1861	TGAACTCATGAAAGAGAGGCGCATTTACTTCAACTGTGTCAACTGCAGCAATGCAGAGAGGAA	1920
Db	1976	TGAATCATGAAAGAGAAAGGCACTTACTTCAACTGTGTCAACTGTGCACAAATGACACACAGGAA	2035
QY	1921	TGAATTTGAGTTAGAAATTTGCCACTGTGTGATTCGCAAAAGTAAGTAAGTATCCCTTGGAAAT	1980
Db	2036	TGAATTTGAGTTAGAAATTTGCCACTGTGTGATTCGCAAAAGTAAGTAAGTATCCCTTGGAAAT	2095
QY	1981	GTCGCCAAAAAGATTCAGGCTGCAGTTTAATTAATAAAGAAAGAACTGCGAGAGGTATCA	2040
Db	2096	GCTCTCAAAATGATTCAAATTCACATCTCAATTAAGAAAGAAAGATCACTCGTGGAGGTCTCG	2155
QY	2041	TGCACCAAGGCGCAGACAGACAGAAAGCTTGTCACAAAAGAGACTGAATGATGATGTATCC	2100
Db	2156	TGCATCCAAAGGCCAAGACAGAAAGCTTAATACAAAAGGCTGTGATGAAGTAATAC	2215
QY	2101	TCCAGTTTCCTCTGTGAGAGATTCGAGAGCTGGAACCTCACTAATAGGCGCTTAATTTGTGTG	2160
Db	2216	TCCAGTTTCCTCTGTGAGAGATTCGAGAGCTTAATTAAGTAATAGGCGCTTAATTTGTGTG	2275
QY	2161	TGGAATAATTTTGCTATTTAATAAGSAGGCGCGACACACATTCCTCAATTAATTTTC	2220
Db	2276	TGCTGTATTTTGTGCTATTTAATAATGAGGCGCTGCACACACATTCCTCAATTAATTTTC	2335
QY	2281	CAGTCTTCGTAATGTTTCTAGTCCCTTGGAAATATTTCTTTTATTCATTTTCCCTCA	2340
Db	2396	CTTGTTTTCAGTATTTGTTTAGCCCTTGGAATATTTCTTTTATTCATTTTCCCTCA	2455
QY	2341	GGGCTTCACATTTGGCAAAAGCTGGGAGATCCCTCAATGAGGCGTTCGATACATGGTTTT	2400
Db	2456	GGGTTTCACATTTGGCAAAAGCTGGAGATCCCTCAACAGGCGCTCCGATACATGGTTTT	2515
QY	2401	CAGATCCATGCTGAGACAGGATGTCAATGTGTTGATACCTTAATAAACCACTGGAGAC	2460

Db	3536	CATACATGCTTCATCGAGTCACTGGCTAAATAATATAGACATAAAGTACGACAAAGG	3655
QY	3541	AACCCAGCTCTCTGGTGGCCAGAAACAGGCAATTCGCATAGCTGCGCTCTTTAGACA	3600
Db	3656	AACACAGCTCTCTGGTGGCCAGAAACAGGCAATTCGCATAGCTGCGCTCTTTAGACA	3715
QY	3601	GCCTCATTTTTTGGTTTTGGATGAACTACATCAGCTCTGCATACAGAAAGTGAAGAGT	3660
Db	3716	GCCTCATTTTTTGGTTTTGGATGAACTACATCAGCTCTGCATACAGAAAGTGAAGAGT	3775
QY	3661	TGTCCAGAAAGCCCTGGCAACAAGCCAGAAAGGCCACCTCTCATTTGATGCGCCACCG	3720
Db	3776	TGTCCAGAAAGCCCTGGCAACAAGCCAGAAAGGCCACCTCTCATTTGATGCGCCACCG	3835
QY	3721	CTTGTCACCAATCCAGAAATGCAGATTTATATAGTGGTGTTCAGATGTCGAATGCAAGTCAAGA	3780
Db	3836	CCGTGCACCATCCAGAAATGCAGATTTATATAGTGGTGTTCAGATGTCGAATGCAAGTCAAGA	3895
QY	3781	GCATGGCACAATCAACAGCTGCTGGCCCGCAAAAAGGATATCTTTTCCATGTCAGTGT	3840
Db	3896	GCATGGCACAATCAACAGCTGCTGGCCCGCAAAAAGGATATCTTTTCCATGTCAGTGT	3955
QY	3841	CCAGGTGGAGAAAGCCCTGAGTGAAGTGGCCATATGAGCTGTAAATATTTTTTAT	3900
Db	3956	CCAGGTGGAGAAAGCCCTGAGTGAAGTGGCCATATGAGCTGTAAATATTTTTTAT	4015
QY	3901	ATTGTGTTAAACATGCAATTTATTCAAAGTTTAAAGGTGAGCAGCTTACTGCAAAACT	3960
Db	4016	ATT--TGTTAGATATGACATTTATTCAAAGTTAA-----AGCAAACTTTACAGAAAT	4068
QY	3961	ATGAGAAACATGCTGTTTAACTTTCTGTCGCAACGAGATCATTTCCACCAAGTTCAAG	4020
Db	4069	ATGAGAAACATGCTGTTTAACTTTCTGTCGCAACGAGATCATTTCCACCAAGTTCAAG	4116
QY	4021	AGCTTCAGATTTTATATTAATAAAGAACCA--AAGAAACATTTCTGATGAGATTAAT	4078
Db	4117	----AGGACCTCGTAATTAAGAAACGAGTGAAGACATCATCAAGTGAAGAAAT	4171
QY	4079	ACTGCTGTTAATGCATTAATAAATATAGATTAATCAAGTAGATTTGTTAATAAT	4138
Db	4172	CGTAGTTTAAACGTCATTAATAATTTATTAACGATTAATAAGTATTTTAAAGATPAA	4231
QY	4139	TGTATATATTTTGTATTAATTTT-----ATTGTACTTACGTCGTTCTGTAAGAT	4190
Db	4232	ATGTGTATTTTGTATTAATTTTCCATTTGACGTGTAACGTCGTCCTGCTAAGAT	4291
QY	4191	TATTAAGTGTAAAGTACTG--ATGTTTGAATAAAGTGTAGCTATATATAAAGTAA	4249
Db	4292	TATTAAGTGTAAAGTACTG--ATGTTTGAATAAAGTGTAGCTATATATAAAGTAA	4348
QY	4250	ACTTTATATCAAAAAAAAAAAAAAAAAAAAAA 4279	
Db	4349	ACTTTATATGCTGACTGAAAAAAAAAAAAAAAA 4378	
RESULT 11			
AR091275	AR091275	4669 bp	DNA
LOCUS	Sequence 1	from patent US 594086.	linear
DEFINITION			
ACCESSION	AR091275		
VERSION	AR091275.1	GI:10018030	
KEYWORDS			
SOURCE	Unknown.		
ORGANISM	Unknown.		
REFERENCE	1 (bases 1 to 4669)		
AUTHORS	Mechnitzer, E. and Roninson, I.B.		
TITLE	Methods and reagents for preparing and using immunological agents		
JOURNAL	Specific for P-glycoprotein		
FEATURES	Patent: US 594086-A 1 30-NOV-1999;		
	Location/Qualifiers		
	1..4669		
	/organism="unknown"		

us-09-672-725c-1.rge

Page 25

Query Match	77.9%;	Score 3333.2;	DB 6;	Length 4669;
Best Local Similarity	87.5%;	Pred. No. 0;		
Matches 3753; Conservative	0;	Mismatches 403;		

1007 indels 447 gaps

```
OY      1 GGGAGCGGAGGATCGGATGATCTCGAAGGAGAGGCCCTTAAGGAGTCCAGAGAAGACTT 60
         |||
Db      409 GGAGCGCGAGGTCGGATGATCTTGAAGGGGACCGCAATGAGACAGCAAGAAAGAA 468
OY      61 CTGGAAATGGCGAATAAAAAAGTAAAAAATAGAGAAGAAAGAAAGAACCACTCTGAG 120
         |||
Db      469 CTTTAAAACTGAACAATAAAGGAAAAAGATAGAGAAAGAAAGAACCACTCTGAG 528
OY      121 CAGCTTGAATGTTGCGCTATTCAAAATGGCTTGATAGTGTATATGTGGTGGGAC 180
         |||
Db      529 TGAATTTCAAGTTTCCCTATTCAAAATGGCTTGACAAAGTTGATATATGTGGTGAAC 588
OY      181 AATGGCTGCATCATCATGAGGCTGCACCTCCCTCATGATCTGCTGTTTGGAAACT 240
         |||
Db      589 TTGGCTGCATCATCATGAGGCTGCAGCTTCCCTCATGATCTGCTGTTTGGAAAT 648
OY      241 GACACATAGCTTGGCAATGCGGAATTTCAAGAAACAAACTTTTCCAGTTATATTTAA 300
         |||
Db      649 GACACATAGCTTGGCAATGCGGAATTTGAAGATCTGAGT-----CAACATCAC 702
OY      301 TGAAGTATTACGAACATACACAACTTTATCAACACATCTGGAGAGGAATGACAC 360
         |||
Db      703 TAATGAAGTATATCAATGATACAGGGTCTTCATGATCTGGAGAGACATGACAG 762
OY      361 GTATGCCATATATACAGTGGAGTCGGTCTGGCGTGGCTGCTTATCATCAGGT 420
         |||
Db      763 GTATGCCATATATACAGTGGAGTTGGTCTGGGGTGGCTGCTGCTTATCATCAGGT 822
OY      421 TTCATTGCGTGGCTGCGCAGCAGAGAACAGATACTCAAAATTGAAACAACTTTTTC 480
         |||
Db      823 TTCATTGCGTGGCTGCGCAGCTGGAGAACAAATACAAAAATTGAAACAGTTTTC 882
OY      481 TGCTATCATCGCAGCAGAGATTGGCTGGTTTACGTCATACGTTGGGAGCTTAAAC 540
         |||
Db      883 TGCTATATATCGCAGAGATAGCGTGGTTGATGTGCGCATTTGGGAGGCTTAAAC 942
OY      541 CCGGCTACAGAGAGATGTCCTCAAAATCAATGAAGAAATTGGCAGCAAAATTGGAATGT 600
         |||
Db      943 CCGACTTACAGAGATGTCCTCCAAATTAAAGAGAAATTGGTACAAATTTGGAATGT 1002
OY      601 CTTTACTCAAAAGACAATTTTCAACGGTTTATATGGGGTTTACACTGTGGTGA 660
         |||
Db      1003 CTTTCACTCAATGGCAACATTTTCACTGGTTTATATGGATTTACACTGTGGTGA 1062
OY      661 GCTAACCCCTTGATTTTGGCCATGACCCCTGTTCTTGACATTCAGCCGCATCGGGG 720
         |||
Db      1063 GCTAACCCCTTGATTTTGGCCATGACCCCTGTTCTTGACATTCAGCTGTCTGGGC 1122
OY      721 AAAGATACTATCTTCACTTATGATTAAGAACTTGGCCTATGCAAAAGCTGAGCACT 780
         |||
Db      1123 AAAGATACTATCTTCACTTATGATTAAGAACTTATAGGTATGCAAAAGCTGAGCACT 1182
OY      781 AGCTAATAACTCTTACAGCAATCAGACTGTGATTTGCTTTGGAGCAAAAGAAAG 840
         |||
Db      1183 AGCTAATAAGGTCTTGGCAGCAATTGAACTGTGATTTGATTTGGAGCAAAAGAAAG 1242
OY      841 ACTTGAAGGTACAACAAAATTTAAGAAGTAAAGAAATGGGATTAAGAAAGCTAT 900
         |||
Db      1243 ACTTGAAGGTACAACAAAATTTAAGAAGCTAAAGAAATGGGATTAAGAAAGCTAT 1302
OY      901 CACGCGCAACTTTCAATTTGCTGCGCTTTCTTATGATCTATGCAATCATGCTGCGC 960
         |||
Db      1303 TACAGCCAAATTTCTATATGCTGCTGCTTCTGCTGATCTATGCAATCTATGCTGCGC 1362
OY      961 TTTTGTATGAGACCTCTTGTGCTCTCCAGTAAATTTCTATGAGCAAGTACTCAC 1020
         |||
```

Db	1363	CTTCTGATATGGGAACAACCTTGGTCTCTCCAGGGGAATATCTAATTGGACAAGTACTAC	1422
QY	1021	TGCTCTCTTTCTGTATTAATATGGGGCTTTTATATATGACAGCATGCCAAGCATTTGA	1080
Db	1423	TGTAATCTTTCTGTATTAATATGGGGCTTTTATGTTTGACAGCATCTCCAAAGCATTTGA	1482
QY	1081	AGCAATTTGGCAACGCCAAGAGAGACACTTATGAAATCTTCAAGATTAATTTGCAATTAAC	1140
Db	1483	AGCAATTTGCAAAATGCAAGAGAGACACTTATGAAATCTTCAAGATTAATTTGATTAAGCC	1542
QY	1141	AAGCATTTGACAGTATATTCGAAAGTGGAGATTAACCAATATATTAAGGAAATTTGGA	1200
Db	1543	AAGTATTTGACAGTATATTCGAAAGTGGAGATTAACCAATATATTTAAGGAAATTTGGA	1602
QY	1201	ATTCAAAATATGTTCACTTCACTAGTTACCTCTCTCGAAAGAAATTTAAGATCTTAAAGGCTC	1260
Db	1603	ATTCAAGAAATGTTCACTTCACTAGTTACCTCTCTCGAAAGAAATTTAAGATCTTAAAGGCTC	1662
QY	1261	CAACCTGAAGGTTTCAAGTGGGGCACAACAGTGGCGCTGGTTGGCAAGTGGCTGGCGGAA	1320
Db	1663	GAACCTGAAGGTTTCAAGTGGGGCACAACAGTGGCGCTGGTTGGCAAGTGGCTGGCGGAA	1722
QY	1321	GAGCAGACCGGTGCAGCTGATGTCAGAGGCTTATGACCCCAAGATGGCATGGCTGTAT	1380
Db	1723	GACCAACAAAGTCCAGCTGATGTCAGAGGCTTATGACCCCAAGAGGGATGGCTCAAGTGT	1782
QY	1381	TGATGGACAGGCATTTAGGACCAATTAATGTAAGCATCTTCCGGAAATTAAGTGTGTGT	1440
Db	1783	TGATGGACAGGCATTTAGGACCAATTAATGTAAGGTTTCTACGGGAATTAATGTTGTGTGT	1842
QY	1441	GAGTCAGAGCCCTGTGTGTTTGGCCACAGATAGCTGAAAAATATTCGCTATAGCCCGGA	1500
Db	1843	GAGTCAGAGACCTGTATGTATGTTTGGCCACAGATAGCTGAAAAATTCGCTATAGCCCGGA	1902
QY	1501	AAATCTCAACCTGGTGTGATTTGGAAGCTGTTAAGAAAGCCAAATGCTATGATTTAT	1560
Db	1903	AAATCTCAACCTGGTGTGATTTGGAAGCTGTTAAGAAAGCCAAATGCTATGATTTAT	1962
QY	1561	CATGAACCTACCTAATTAATTTGACACTGTGTTGGAGAGAGGGGCCAGAGTAGAGG	1620
Db	1963	CATGAACCTGCTCATTAATTTGACACCCGTGGTGAAGAGAGGGGCCAGAGTAGAGG	2022
QY	1621	TGGACAGAAACAGGAATGGCCATTCCTGGGCGCTGGTTGGCAACCCCAAGATTTCTTCT	1680
Db	2023	TGGCCACAAGACAGAGATTCGCCATTTGCACAGTCCCTGTTGGCAACCCCAAGATTCCTTCT	2082
QY	1681	GCTGGATGAGGCAACGCTCAGCTCTGGACACTGAAGTGAACAGACGTTGATAGGTCCT	1740
Db	2083	GCTGGATGAGGCAACGCTCAGCTCTGGACACTGAAGTGAACAGACGTTGATAGGTCCT	2142
QY	1741	GGATTAAGGCCAAGAAAGGCCGACACTAATGTGTATGATCTCATGTTGTCTACAGTTG	1800
Db	2143	GGATTAAGGCCAAGAAAGGCCGACCAACCATTTGATATAGCTCATGTTGTCTACAGTTG	2202
QY	1801	TAAATGCCGATGTCAATTCGTGTTTGGATGATGAGAGTCAATTTGGAGAGAAAGAAATATGA	1860
Db	2203	TAAATGCCGATGTCAATTCGTGTTTGGATGATGAGAGTCAATTTGGAGAGAAAGAAATATGA	2262
QY	1861	TGAACCTCAAGAAAGAGAGGCGATTTACTTCAAACTTGTCACAATCCAGACAAGAGAA	1920
Db	2263	TGAACCTCAAGAAAGAGAGGCGATTTACTTCAAACTTGTCACAATCCAGACAAGAGAA	2322
QY	1921	TGAATTTGAATTTGAAGAAATGCACTGGTGAATCCAAAGTGAAGAGTATGCTTGGAAAT	1980
Db	2323	TGAATTTGAATTTGAAGAAATGCACTGGTGAATCCAAAGTGAAGAGTATGCTTGGAAAT	2382
QY	1981	GTCTCCAAAGATTCAGGCTCCAGTTAATTAATAAGAGATCAACTCCGACAGATATACA	2040
Db	2383	GTCTCCAAATATATTCAGAGTCCAGTCAATTAATAAGAAATATCAACTCGTAAGAGATGTCCG	2442
QY	2041	TGCACCAAGGCCCAACAGAAAGCTTGGTATCAAAAGAGGCTTGATAGATGATGCC	2100
Db	2443	TGCATCAACAGGCCCAACAGAAAGCTTGATCAAAAGAGGCTTGGATGATGATGATGCC	2502

Oy	2101	TCGAGTTTCCTTCGAGAGATTCCTGAACTGAACTCAACTGAATGCGCTAATTTGGT	2160
Db	2503	TCGAGTTTCCTTTTGAGAGATTATGAGGCTAAATTTAATCTGAATGGCCCTATTTTGTGT	2562
Oy	2161	TGCTAATATTTTGGTATTATTAAGGAGAGCGCCGCAACCGAGATTTTCCAAATATTTTC	2220
Db	2563	TGGTATATTTTGGCATTTAATATGGAAGGCCCTCGAACCAGATTTGCAATAATATTTTC	2622
Oy	2221	AAGATTTATAGGATCTTTACCCAGATGAGATCCTGAAACAAAGACGATTAAGTAA	2280
Db	2623	AAAGATTATAGGGGTTTTTACAGAAATGTATGATCCTGAAACAAACGACGATTAAGTAA	2682
Oy	2281	CATGTTTTCGTATGTTTCTATAGTCCTGTGGAATATTTTCTTTATTTACATTTTCCCA	2340
Db	2683	CTTTTTCCTATTTCTTTCTTCTAGCCCTTGGAAATATTTCTTTTATTTACATTTTCTTCA	2742
Oy	2341	GGGCTTCATTTTGGCAAGCTGGGGAGATTCCTCACTAAACGGGCTTCGATCATGGTTTT	2400
Db	2743	GGGTTTCATTTTGGCAAGCTGGGAGATCTCTACACAGCGGCTCCGATTAATGGTTTT	2802
Oy	2401	CAGATTCATGCTGAGACAGGATGTGACGTGGTTTGATGACCTTAATAACACATCTGAGC	2460
Db	2803	CCGATTCATGCTGAGACAGGATGTGATGTGTTTATCATGACCTTAATAACACATCTGAGC	2862
Oy	2461	ATTGACACACAGGCTTGCCAAATGATGGCGCTCAAGTTAAAGGGGCTATATGTTCCAGCT	2520
Db	2863	ATTGATCTCCAGGCTGCCCATATATCTCTCAAGTTAAAGGGGCTATATGTTCCAGGCT	2922
Oy	2521	TGCTCTCATTTACCCAGAAATATAGCAATCTTTGGACAGGCAATTATATTCCTTAATCA	2580
Db	2923	TGCTGTAATTTCCAGAAATATATGCAATCTTTGGACAGGAAATATATATCTTCAATCA	2982
Oy	2581	TGGTTGGCAATTAACACTTTTACTCTTTAGCAATTTTACCCATCTTCAATATGACGAGAT	2640
Db	2983	TGGTTGGCACTTAACACTGTTTACTCTTTAGCAATTTGACCATCTTCAATATGACGAGAT	3042
Oy	2641	TGTTGAATGAAATGTTGTCTGAGACACACCTGAAAGATTAAGAAAGCTATAGAGAC	2700
Db	3043	TGTTGAAATGAAATGTTTGTGTGACACAGACCTGAAAGATTAAGAAAGACTATAGAGGTC	3102
Oy	2701	TGGGAAGATTTGCTATAGAAAGCACTCGAAAACTTCCGAACTGTTGTTCTTTGACTGGGA	2760
Db	3103	TGGGAAGATCTGCTACTGAAGCAATAGAAAACTTCCGAAACCGTTGTTTGTGACTCGGA	3162
Oy	2761	GCAGAGTTTGAAATACATGATGACACAGATTTGGCAATGACATACAGAAACTCTTTGAG	2820
Db	3163	GCAGAGTTTGAAATATGATGATGCTCGAGATTTGGCAGTACATATCAGAAACTCTTTGAG	3222
Oy	2821	GAAAGCACACATTTCCGGGCTCCTATTTTCTATCCACGACAGCAATGATGATTTTCTCA	2880
Db	3223	GAAAGCACACATTTTGGAAATTAATTTCTCTTCCACCCAGGCAATGATGATTTTCTCA	3282
Oy	2881	TGCTGGCTGTTTCCGTTTGGTGCCCTACTTTGGGCAAAATGAGTTTCAATGAACTTTCAGA	2940
Db	3283	TGCTGGATGTTTCCGTTTGGAGCCCTACTTGTGTGGCAATTAATCTATGAGCTTTGAGGA	3342
Oy	2941	TGTTCTTTTGGTATTTCTCAGCTAATTTGCTTTTGTGTCATGGCAGTGGGGCAGGTAGTTC	3000
Db	3343	TGTTCTTTTGGTATTTTCAAGCTGTTGTGTGTGGGCCATGGCCGTGGGCAAGTATGTTTC	3402
Oy	3001	ATTTGCTCTGACATGAGCCAAAGCAATACAGACAGCCACAGTCATCATGATCATTTGA	3060
Db	3403	ATTTGCTCTGACATGAGCCAAATATACAGACAGCCACATCATGATCATTTGATGA	3462
Oy	3061	AAAAAGCCCTTGATTTGACAGCTTACAGCCCTTACGGCTCTAAGGCAAAATACGTGGAGG	3120
Db	3463	AAAAAGCCCTTTGATTTGACAGCTTACAGCGGAGAGGCTTAATGCGCAACATTTGGAAG	3522
Oy	3121	AAATGACATTTATAGAGTGTGTTCACATTTCCACTGACACAGACATCCCGTGT	3180
Db	3523	AAATGACATTTTGTGTAATGTATATCAACATATCCACCCGACGACATCCCAATGCT	3582

QY	3181	CCAGGGGCTGAGCCCTCGAGGTGAAGAGGGCCAGACGCTGGCCCTCTGATGATGACGATGGG	3240
Db	3583	TCAGGCACTGAGCGTGGAGGTGAAGAGGGCCAGACGCTGGCCCTCTGATGATGACGATGGG	3642
QY	3241	CTGTGGAGAGCACAGTTGTTCAAGCTCTTGAAGCCGCTTCTATGACCCCTTGGCTGTTCC	3300
Db	3643	CTGTGGAGAGCACAGTTGTTCAAGCTCTTGAAGCCGCTTCTATGACCCCTTGGCTGTTCC	3702
QY	3301	AGTCTTAATTGATGACGAAGATTAAGACACCTTAATGTCCTCAATGGCTCCGAGCACACT	3360
Db	3703	AGTCTCTTGTGATGGCAAGAAATTAACGACTGAATGTTCAATGTTGAGTCTCCGAGCACACT	3762
QY	3361	GGGCACTGTCTCTCAGAGACCCATCTGTTGACTGACGAGATTGCCAGAACTTGGCTTA	3420
Db	3763	GGGCACTGTCTCTCAGAGACCCATCTGTTGACTGACGAGATTGCCAGAACTTGGCTTA	3822
QY	3421	TGGAGACAAACGCCGGGTCTATCAGATGAGAGATTATGCAAGCCAGCCAGAGAGGCCAA	3480
Db	3823	TGGAGACAAACGCCGGGTCTATCAGATGAGAGATTATGCAAGCCAGCCAGAGAGGCCAA	3882
QY	3481	CATTACCACTTCAATGCAACACTCCCTGAGAAATACAAACCCAGACTATGAGACAAAGG	3540
Db	3883	CATTACCTCTTCAATGCAAGCTCTGCTTAATTAATATGCACTAAAGTAGACACAAAGG	3942
QY	3541	AACCCAGCTCTCTGATGGCCAGAAAGCCGCAATTGGCAATGACTGCGCTCTGTTGACA	3600
Db	3943	AATCCAGCTCTCTGATGGCCAGAAAGCCGCAATTGGCAATGACTGCGCTCTGTTGACA	4002
QY	3601	GCCTCATTTTTCCTTTTGGATGAGCTATCATAGCTGCTGATACAGAAATGAGAAAGT	3660
Db	4003	GCCTCATTTTTCCTTTTGGATGAGCTATCATAGCTGCTGATACAGAAATGAGAAAGT	4062
QY	3661	TGTCCAAAGAGCCCTGGACAAAGCCAGAGAGGCCACCCTGCTGATGATGGCCACCG	3720
Db	4063	TGTCCAAAGAGCCCTGGACAAAGCCAGAGAGGCCACCCTGCTGATGATGATGCTCACCG	4122
QY	3721	CTTGTCCACCATTCAGAAATTCAGATTAATAGTGGTGTTCAGAAATGGCAAGTCAAGGA	3780
Db	4123	CTTGTCCACCATTCAGAAATTCAGATTAATAGTGGTGTTCAGAAATGGCAAGTCAAGGA	4182
QY	3781	GCATGGCACATCAATCAACAGCTCTGGCCAGAAAGCATATTTTCCATGGTCAGAGT	3840
Db	4183	GCATGGCACATCAATCAACAGCTCTGGCCAGAAAGCATATTTTCCATGGTCAGAGT	4242
QY	3841	CCAGGCTGGAGCAAAACCGCTAGTGAACGTGGCCATGAGCTGTAAATATTTTAAAT	3900
Db	4243	CCAGGCTGGAGCAAAACCGCGAGGAACTCTGACGTGATGAGATGTTAAATACCTTTTAAAT	4302
QY	3901	AATTGTGTTAAACATGAGCATTTTAATCAAAAGTTAAAGGTGAGCACTCTCGAAAAACT	3960
Db	4303	AATTTTGTGTTAAACATGAGCATTTTAATCAAAAGTTAAAGGTGAGCACTCTCGAAAAACT	4358
QY	3961	AATGTAAGAACTACCTGTTTAAATTTCTTGCTGCACATGAGATATTCACCAAGTTCCG	4020
Db	4356	AATGTAAGAACTACCTGTTTAAATTTCTTGCTGCACATGAGATATTCACCAAGTTCCG	4402
QY	4021	AATCTTCGATTTTATTAATTAAGAACCA--AAGAAACATTAATTCGATGGAATTAAT	4078
Db	4404	AATCTTCGATTTTATTAATTAAGAACCA--AAGAAACATTAATTCGATGGAATTAAT	4450
QY	4079	ACTGGTGTAAATTCGATTAATTAATTAATTAAGAAATTAAGTAAATGTTGTTTAATTAAT	4133
Db	4459	ACTGGTGTAAATTCGATTAATTAATTAATTAAGAAATTAAGTAAATGTTGTTTAATTAAT	4511
QY	4139	TGTATTAATTTTGTATATATTTT-----ATTGTAACCTTACTGCTTCTCAAGAAAT	4199
Db	4519	TGTATTAATTTTGTATATATTTT-----ATTGTAACCTTACTGCTTCTCAAGAAAT	4571
QY	4191	TATTAAGAAAGTGAATAAGTACTG--AAGTGTGAATTAAGCTACTGATTAATTAATACTAA	4244
Db	4579	TATTAAGAAAGTGAATAAGTACTGATTAAGTGTGAATTAAGCTACTGATTAATTAATACTAA	4633
QY	4250	ACTTTTATATCAAAAAAAAAAAAAAAAAAAAAA 4279	

Db 4636 ACTTTCATGTGAAAAA 4665

RESULT 12
AR203322 4669 bp DNA linear PAT 20-JUN-2002

DEFINITION Sequence 1 from patent US 6365357.
ACCESSION AR203322
VERSION AR203322.1 GI:21499681
KEYWORDS
SOURCE Unknown.
ORGANISM Unclassified.
REFERENCE 1 (bases 1 to 4669)
AUTHORS Mechtner,E. and Fruehauf,J.
TITLE Methods and reagents for preparing and using immunological agents specific for P-glycoprotein
JOURNAL Patent: US 6365357-A 1 02-APR-2002;
FEATURES Location/Qualifiers
source 1..4669
BASE COUNT 1393 a 894 c 1130 g 1252 t
ORIGIN

Query Match 77.9%; Score 333.2; DB 6; Length 4669;
Best Local Similarity 87.5%; Pred. No. 0;
Matches 3753; Conservative 0; Mismatches 493; Indels 44; Gaps 8;

QY 1 GGAGCCGAGGTGCGGATGATCCGGAAGAGAGCGCTAAGGGAGTGCAGAGAACAATT 60
Db 409 GGAGCCGAGGTGCGGATGATCCGGAAGAGAGCGCTAAGGGAGTGCAGAGAACAATT 60
QY 61 CTGGAATAGGGCAAAAAAGTAAAAATGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 120
Db 469 CTTTCTTTAACTGAACAATAAAGTGAAGAGAGAGAGAGAGAGAGAGAGAGAGAG 120
QY 121 CAGCTTGGCAATGTTCCCTATTCAAATGCGTTGATAGTTTATATGTTGGTGGGAC 180
Db 529 TGTATTTCAATGTTTGGCTATTCAAATGCGTTGATAGTTTATATGTTGGTGGGAC 180
QY 181 AATGGCTGCCATCCATCCATGAGCTGCACCTCTCATGATGCTGTTTGGAAACAT 240
Db 589 TTGGCTGCCATCCATCCATGAGCTGCACCTCTCATGATGCTGTTTGGAAACAT 240
QY 241 GACAGATAGCTTTGCAATGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 300
Db 649 GACAGATATCTTGCATATGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 300
QY 301 TGAAGATATGAG 360
Db 703 TATATGAGAGATATGAG 360
QY 361 GTATGCTATTTATTTACAGTGGATGCGTGGCTGGCTGGCTGGCTGGCTGGCTGGCT 420
Db 763 GTATGCTATTTATTTACAGTGGATGCGTGGCTGGCTGGCTGGCTGGCTGGCTGGCT 420
QY 421 TTCAATTCGCTGGCTGGCTGGCTGGCTGGCTGGCTGGCTGGCTGGCTGGCTGGCTGG 480
Db 823 TTCAATTCGCTGGCTGGCTGGCTGGCTGGCTGGCTGGCTGGCTGGCTGGCTGGCTGG 480
QY 481 TGTATCATGCGACAGAGAGATTTGGCTGGCTGGCTGGCTGGCTGGCTGGCTGGCTGG 540
Db 883 TGTATCATGCGACAGAGAGATTTGGCTGGCTGGCTGGCTGGCTGGCTGGCTGGCTGG 540
QY 541 CCGGCTCAAGAGAGATTTGCCAAATCAATGAAGAGATTTGGCAAGAGATTTGGATGTT 600
Db 943 CCGACTTAAAGATGATGTTCCAAAGATTAATGAAGAGATTTGGCAAGAGATTTGGATGTT 600
QY 601 CTTTCACTCAATAGCAATTTTACCGGTTTATATGTTGGGTTTAAACAAGTGGTTGAA 660
Db 1003 CTTTCACTCAATAGCAATTTTACCGGTTTATATGTTGGGTTTAAACAAGTGGTTGAA 660

QY 661 GCTAACCTTTGATATTTTGGCCATGACGCCCTGTTCTTGAGCTTTCAGCCCATCTGGGC 720
Db 1063 GCTAACCTTTGATATTTTGGCCATGACGCCCTGTTCTTGAGCTTTCAGCTGCTGTTGGGC 720
QY 721 AAGATATATCTTCAATTTACGATTAAGAACTCTTGGCCATGACAAAGCTGGAGAGT 780
Db 1123 AAGATATATCTTCAATTTACGATTAAGAACTCTTGGCCATGACAAAGCTGGAGAGT 780
QY 781 AGCTGAAGAGTCTTAGCAATCAGAACTGATTTGCTTGGAGAGAGAGAGAGAGAG 840
Db 1183 AGCTGAAGAGTCTTAGCAATCAGAACTGATTTGCTTGGAGAGAGAGAGAGAGAG 840
QY 841 ACTTGAAGATCAACAAAAATTTAGAAAGCTTAAAGATTTGGATTAAGAAAGCTAT 900
Db 1243 ACTTGAAGATCAACAAAAATTTAGAAAGCTTAAAGATTTGGATTAAGAAAGCTAT 900
QY 901 CAGGCCCAACATTTCTATTTGCTGGCTTCTTATGATATGATATGATATGATATGAT 960
Db 1303 TACACCCAAATTTCTATTTAGGCTGCTTCCGTCGATCTATGATATGATATGATATG 960
QY 961 TTTCTGATGAGGACCTCTTGGCTCTCTCACTGAATATTTATTTGACAAAGTACTCAC 1020
Db 1363 CTTCTGATGAGGACCTCTTGGCTCTCTCACTGAATATTTATTTGACAAAGTACTCAC 1020
QY 1021 TGTCTTCTTCTGATATTTATTTGAGGCTTTTATGATTTGAGAGAGAGAGAGAGAG 1080
Db 1423 TGTATCTTCTTCTGATATTTATTTGAGGCTTTTATGATTTGAGAGAGAGAGAGAG 1080
QY 1081 AGCTTTGCAAG 1140
Db 1483 AGCTTTGCAAG 1140
QY 1141 AAGCATTGACAGTATTTGCAAGAGTGAATACCATATATTTAAGGAAATTTTGA 1200
Db 1543 AAGTATTTGACAGTATTTGCAAGAGTGAATACCATATATTTAAGGAAATTTTGA 1200
QY 1201 ATTCAAAATATTTCACTTATTTACCTTCTCTGAAAAGAGATTTAAGATTTTAAAG 1602
Db 1603 ATTCAAAATATTTCACTTATTTACCTTCTCTGAAAAGAGATTTAAGATTTTAAAG 1602
QY 1261 CAACCTGAAGGTTTCAGAGTGGGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1320
Db 1663 GAACCTGAAGGTTTCAGAGTGGGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1320
QY 1321 GAGCAGAGAGGCTGCTGATGATGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1722
Db 1723 GAGCAGAGAGGCTGCTGATGATGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1722
QY 1381 TGATGAGAGAGATTTAGAGCATTAATTTAAGAGCATTTTGGGAAATTTACGTGTGT 1440
Db 1783 TGATGAGAGAGATTTAGAGCATTAATTTAAGAGCATTTTGGGAAATTTACGTGTGT 1440
QY 1441 GAGTGAAGAGGCTGCTGATGATGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1842
Db 1843 GAGTGAAGAGGCTGCTGATGATGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1842
QY 1501 AATGTCACTGATGATGATTTGAGAAAGCTTTAAGAGAGAGAGAGAGAGAGAGAGAG 1902
Db 1903 AATGTCACTGATGATTTAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1902
QY 1561 CATGAATACCTTAATTAATTTGACACTGCTGTTGAGAGAGAGAGAGAGAGAGAGAG 1962
Db 1963 CATGAATACCTTAATTAATTTGACACTGCTGTTGAGAGAGAGAGAGAGAGAGAGAG 1962
QY 1621 TGGACAGAGAGAGAGATTTGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 2022
Db 2023 TGGACAGAGAGAGAGATTTGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 2022
QY 1681 GCTGATGAGGAG 1740
Db 2083 GCTGATGAGGAG 1740
QY 1741 GGATTAAGGAG 1800

|||||
Db 2143 GGATAGGCGGAAAAGGTGGACCACTTGTGATGCTCATGCTTTGTCTACAGTTGG 2202
QY 1801 TATGCCGATGTCATTTGCTGTTTGTATGATGAGTCATTTGTGGAGAAATATGTA 1860
Db 2203 TATGCTGACGTCATCGCTGGTTTCGATGATGAGTCATTTGTGGAGAAATATGTA 2262
QY 1861 TGAACATGAAAGAAAGGGGATTTACTTCAAACTTGTGCACAAATGACAGACAGGAAA 1920
Db 2263 TGAACATGAAAGAAAGGGGATTTACTTCAAACTTGTGCACAAATGACAGACAGGAAA 2322
QY 1921 TGAATTTGAGTTAGAAAATGCGCATGTCGATCCAAAAGTGAAGTATGCTTTGAAAT 1980
Db 2232 TGAAGTTGAAATTTAGAAAATGCGCATGATGAATCCAAAAGTGAAGTATGCTTTGAAAT 2382
QY 1981 GTCTCCAAAAGATTCAGGGTCCAGTTTAAATAAAGAGATCAATCGCAGAGTATACA 2040
Db 2383 GTCTTCAAAATGATTCAGATCCAGTCTAATTAAGAAAAAGATCAATCGTAGAGTGTCCG 2442
QY 2041 TGCACCAAGGCCAAGACAGAAAAGCTTGTGTCAAAAGAGACTTGAATGGAATGATAC 2100
Db 2443 TGGATACAAAGCCCAACACAGAAAGCTTAGTACCAAGAGGCTGTGATGAAATGATAC 2502
QY 2101 TCCAGTTTCTCTGAGAGATCTGAGAGCTGAAGCTCAACTCAATGAAATGCTTATTTGTGT 2160
Db 2503 TCCAGTTTCTCTGAGAGATTTAGAAAGCTAATTTAACTGAATGAGCTTATTTGTGT 2562
QY 2161 TGGATATTTTGTCTATTTAAACGAGGCTGCAACAGATTTCAATTAATTTTC 2220
Db 2563 TGGTATTTTGTGCTATTTAAATGAGAGGCTGCAACAGATTTCAATTAATTTTC 2622
QY 2221 AAGGATTTAGGAGTCTTTACCCGAGATGAGATCTGAAACAAACGACAGATATGTA 2280
Db 2623 AAGGATTTAGGAGTCTTTACCAAAATGATGATCCGAAACAAACGACAGATATGTA 2682
QY 2281 CATGTTTCTGATTTGTTTCTAGTCTTGGATATTTCTTTATTAATTAATTTTCTTCA 2340
Db 2683 CTGTTTCTGATTTGTTTCTAGGCTTGGAAATTTCTTTATTAATTTTCTTCA 2742
QY 2341 GGGCTTCACATTTGGCAAGGCTGGGAGATCTCACTAAGCGGCTTGCATATGTTT 2400
Db 2743 GGGCTTCACATTTGGCAAGGCTGGGAGATCTCACTAAGCGGCTTGCATATGTTT 2802
QY 2401 CAGATCCATGCTGAGACAGAGATGTCAGTGTGATGACCCCTAAACCACTGAGAG 2460
Db 2803 CAGATCCATGCTGAGACAGAGATGTCAGTGTGATGACCCCTAAACCACTGAGAG 2862
QY 2461 ATTGAACAACAGGCTTGGCAATGATGCGCTCAAGTTAAAGGGCTATAGTTCCAGGCT 2520
Db 2863 ATTGAACAACAGGCTTGGCAATGATGCGCTCAAGTTAAAGGGCTATAGTTCCAGGCT 2922
QY 2521 TGGTGTATTTACCCAGAAATATGACAAATCTTGGGACAGGCAATTTATCTTAATCTA 2580
Db 2923 TGGTGTATTTACCCAGAAATATGACAAATCTTGGGACAGGCAATTTATCTTAATCTA 2982
QY 2581 TGGTGTATTTACCCAGAAATGTTGCTGAGACAGACTGAAAGATAGAAAGGCTAGAAAGAG 2640
Db 2983 TGGTGTATTTACCCAGAAATGTTGCTGAGACAGACTGAAAGATAGAAAGGCTAGAAAGAG 3042
QY 2641 TGTGAAATGAAATGTTGCTGAGACAGACTGAAAGATAGAAAGGCTAGAAAGAG 2700
Db 3043 TGTGAAATGAAATGTTGCTGAGACAGACTGAAAGATAGAAAGGCTAGAAAGAG 3102
QY 2701 TGGGAGATTTCTACAGAAAGCAATCGAAATCTCGAAATGTTGTTCTTTGACTCGGGA 2760
Db 3103 TGGGAGATTTCTACAGAAAGCAATCGAAATCTCGAAATGTTGTTCTTTGACTCGGGA 3162
QY 2761 GCGAAGATTTGAATATGATGACAGAGTTGCAAGTACATACAGAAATCTTTGAG 2820
Db 3163 GCGAAGATTTGAATATGATGATGCTCAGAGTTGCAAGTACATACAGAAATCTTTGAG 3222
QY 2821 GAAAGCACATCTTCTGGGGCTCATTTTCTATCACCCAGCATGATGATTTTCTTA 2880
|||||

Db 3223 GAAAGCACATCTTCTGGAGATTTACATTTTCTTCCCTTCACCCAGGCAATGATGATTTTCTTA 3282
QY 2881 TGGTGTATTTCCGGTTTGTGCTTACTGCTGAGCAATGATGATGATGATGAGCA 2940
Db 3283 TGGTGTATTTCCGGTTTGTGAGCTTACTGCTGAGCAATGATGATGATGAGCA 3342
QY 2941 TGTCTTTTGGATTTCTGAGCTATGCTTTGGTGGTCCATGTCAGTGGGAGGCTGACTTC 3000
Db 3343 TGTCTTTTGGATTTCTGAGCTATGCTTTGGTGGTCCATGTCAGTGGGAGGCTGACTTC 3402
QY 3001 ATTGCTCTGATGATGCAAAAGCCAAAGTATGACAGCCCACTGATCATGATCATTTGA 3060
Db 3403 ATTGCTCTGATGATGCAAAAGCCAAAGTATGACAGCCCACTGATCATGATCATTTGA 3462
QY 3061 AAAAGCCCTCTGATGATGACAGCTGACAGCCCTGACAGCCCACTGATCATGATCATTTGA 3120
Db 3463 AAAAGCCCTCTGATGATGACAGCTGACAGCCCTGACAGCCCACTGATCATGATCATTTGA 3522
QY 3121 AAATGTGATTTTAAATGAGTCTGTTCAATCTATCCACTGACCAATCCCGTCT 3180
Db 3523 AAATGTGATTTTAAATGAGTCTGTTCAATCTATCCACTGACCAATCCCGTCT 3582
QY 3181 CCAGGCTGAGCTCCAGAGTGAAGAGGCGAGAGCTGAGCTCTGATGAGCAATG 3240
Db 3583 CCAGGCTGAGCTCCAGAGTGAAGAGGCGAGAGCTGAGCTCTGATGAGCAATG 3642
QY 3241 CTGTGGAAAGACAGATTTGCTTCACTCTGAGAGCTCTTATGACCCCTTGGCTGTTTC 3300
Db 3643 CTGTGGAAAGACAGATTTGCTTCACTCTGAGAGCTCTTATGAGCCCTTGGCAGGAA 3702
QY 3301 AGTGTATTTGATGCAAAAGATTAAGCACTGATGATGCTGATGATGATGATGATGAT 3360
Db 3703 AGTGTATTTGATGCAAAAGATTAAGCACTGATGATGATGATGATGATGATGATGAT 3762
QY 3361 GGGCATGCTGTCTGAGAGGCTATCTGTTGATGATGATGATGATGATGATGATGAT 3420
Db 3763 GGGCATGCTGTCTGAGAGGCTATCTGTTGATGATGATGATGATGATGATGATGAT 3822
QY 3421 TGGAGACAAAGCGGCTGCTATCAATGAAAGATTAATGATGATGATGATGATGATGAT 3480
Db 3823 TGGAGACAAAGCGGCTGCTATCAATGAAAGATTAATGATGATGATGATGATGATGAT 3882
QY 3481 CATACACACTTCTATGAGACACTCTCTGATGATGATGATGATGATGATGATGATGAT 3540
Db 3883 CATACACACTTCTATGAGACACTCTCTGATGATGATGATGATGATGATGATGATGAT 3942
QY 3541 AACCCAGCTCTGCTGAGGCGCAAGACAGGCAATGCTGATGATGATGATGATGATGAT 3600
Db 3943 AACCCAGCTCTGCTGAGGCGCAAGACAGGCAATGCTGATGATGATGATGATGATGAT 4002
QY 3601 GCTCATATTTTGTGTTTGGATGAAAGCTACATGCTGCTGATGATGATGATGATGATGAT 3660
Db 4003 GCTCATATTTTGTGTTTGGATGAAAGCTACATGCTGCTGATGATGATGATGATGATGAT 4062
QY 3661 TGTCCAAAGAGCCCTGAGCAAAAGCCAGAGAGGCGCACCTGATGATGATGATGATGAT 3720
Db 4063 TGTCCAAAGAGCCCTGAGCAAAAGCCAGAGAGGCGCACCTGATGATGATGATGATGAT 4122
QY 3721 CTGTGCAACATCCAGATGAGATTAATGATGATGATGATGATGATGATGATGATGAT 3780
Db 4123 CTGTGCAACATCCAGATGAGATTAATGATGATGATGATGATGATGATGATGATGAT 4182
QY 3781 GCATGGACATCAACAGCTGCTGAGGCGCAAGAGGCGCATATTTTCAATGATGATGAT 3840
Db 4183 GCATGGACATCAACAGCTGCTGAGGCGCAAGAGGCGCATATTTTCAATGATGATGAT 4242
QY 3841 CCAGGCTGAGCAAGGCGCTGATGATGATGATGATGATGATGATGATGATGATGAT 3900
Db 4243 CCAGGCTGAGCAAGGCGCTGATGATGATGATGATGATGATGATGATGATGATGAT 4302
QY 3901 ATTGTTTAAACATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 3960
Db 4303 ATT - TGTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 4355
|||||

RESULT 13				
LOCUS	AX391099			
DEFINITION	AX391099	4646 bp	DNA	linear
ACCESSION	AX391099	Sequence 1 from Patent WO0210205.		PAT 19-MAR-2002
VERSION	AX391099.1	GI:19584228		
KEYWORDS				
SOURCE				
ORGANISM	human.			
	Homo sapiens			
	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;			
	Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.			
REFERENCE				
AUTHORS	Ruth,A. and Roninson,I.			
TITLE	Mutations of the mdrl P-glycoprotein that improve its ability to			
	confer resistance to chemotherapeutic drugs			
JOURNAL	Patent; WO 0210205-A1 07-FEB-2002;			
	THE BOARD OF TRUSTEES OF THE UNIVERSITY OF ILLINOIS (US)			
FEATURES				
source	location/Qualifiers			
	1..4646			

425. .4267
/note="unnamed protein product"
/protein_id="CAD28599.1"
/db_xref="GI:19584229"
/translation="MDLRGDRNGMKKKKKKPKLNKSEKKKKKKPKPVSVFSMFYSN
WDKLVKVVGLTALITTHGAGLPLMLVFGKDTLTFAAAGLIEDLMKSNITRSDITDG
FPMNLDEDMRTARYYITSGIGAGVLAAVITDSEFWCLAGAGLQIKIRKQFFHAIMQEI
LAISHPDLGVELNTRLITDVSRIINFGIDKIGMFQSVATPEETFGYGFGRKKLTLTV
YKNNLEAKRIGIKKAITANISIGAAFLTAAGAGVAEATLALRYTIAFGGKKTELER
FESVLIGAFSVGQASPSIEAFANRGAAYEETPIIDKRPISDSYSGSKRPNTIGV
ERNRHHSPKSRKEVKILKGLNLTAVSGOAVLVAEYVATYIAERIKYGRNVAETDEAVAN
VAVSDQDITITNVRFLRELIGVSOSEPVETATYIAERIKYGRNVAETDEAVAN
AVDEIMKRLPKHKEDTLVGRGKQKSGGOKQIAIALKALVRNPKLTLLDENTSALDTESE
AVOVALAKARKGRPTTYIYARLSTYANRVAIVAGFDVGYIVKGNHDELMKRGYFK
LVTOTACNNEVELEMADEKSEIDALEMSSNDSRSLIRKSTRSVRSGQAQPKL
SKPKELDDSIPEVSPFKRMKNTLTEMPEYVAVGCAIINGGLQPAFAPIFSKIDGRL
RIDDEIRIRQNSNLSFLTEALGIIISFTFFGLFFFGKAGETLIRKLRIYWRSMRLR
QDVSEFDEPKNTGALTTLRLANDAAQYKAGIAGSLAVITQNTANLGTGIIISITGMO
KFEHHYASQILOVYRNSLRKRAHIFGTFEFTPTOMAYEVSAGCGRGATLVAAKLMSE
DYLVTEASVVGAGAAVGSFPAADYAKATISAHITATYIETKPLIDISTSTEGLMEDP
LEGANTFEEVEYENPTRPDQIVGLGYSLEKPKQTLALVGGSGCKSTVYOLBERPD
PLAGVLTDEVEIKRLRNQVLRHLGIVSGHGLVSGDQKORIAIAVALRYGRHIIILDEATS
RAAREANIHAETESLPKISTKVGDDGTOLSSGOKORIAIAVALRYGRHIIILDEATS

ALDPESKVEQALDKAREGRICVIAHRLSTIONADLIVFONGRVEKHTHOOLLA									
OKGIFESMVSVOGTRO"									
BASE COUNT	1372 a	892 c	1130 g	1252 t					
ORIGIN									
QY	1	GGAGCGCCAGGT	CGGAGTGCATCCTG	AAGAGAGGCCGCT	TAAGGGGAGTGCAGGAACACTT	60	77.8%;	Score 3327.6;	DB 6;
								Best Local Similarity 87.6%;	Length 4646;
Db	409	GGAGCGCCAGGT	CGGAGTGCATCCTG	TAAGGGGAGGCCCACT	TGAGGAGGAGCAAGAGAGAA	468		Prod. No. 0;	
								Mismatches 484;	Indels 44;
QY	61	CTGGAAAAATGGGCA	AAAAAAGTAAAAAATG	GAAGAAAGAAAGAAACCA	CTGTCAG	120		Gaps 8;	
Db	469	CTTTTAACTGAAACA	ATTAAGTGAAGAAAGTAA	GAAGAAAGAAACCACTG	TCAG	528			
QY	121	CACGTTTGCAATGTT	GGCTATTCAAATGGCT	TGATAGTGTATGTTATG	TGTTGGGGAC	180			
Db	529	TGATATTTCAATGTT	GGCTATTCAAATGGCT	TGCAAGTTGATATGTTGG	GGGAGAC	588			
QY	181	AATGGCTGCATCATC	CCATGAGCTCCCTCTC	ATCATGATGCTGCTTTT	TGGAAACAT	240			
Db	589	TTTGCGTCATCATC	ATCATGAGCTCCCTCTC	ATCATGATGCTGCTTTT	TGGAAACAT	648			
QY	241	GACGATATGCTTTG	CAAAATGCAGGAATTT	CAAGAAACAAACTTTT	CCAGTTAAATTA	300			
Db	649	GACGATATGCTTTG	CAAAATGCAGGAATTT	CAAGAAACAAACTTTT	CCAGTTAAATTA	702			
QY	301	TGAAGATATTAACA	ATAACAAATTTTCTAT	CACCATTTGGAGAGAAAT	TGACAC	360			
Db	703	TAAATGAAGATATTA	ATAACAAATTTTCTAT	CACCATTTGGAGAGAAAT	TGACAC	762			
QY	361	GATGCTATTTATTA	CAGTGGATGGTGGCTG	GGCTGGTGGCTCTTACAT	TCACGT	420			
Db	763	GATGCTATTTATTA	CAGTGGATGGTGGCTG	GGCTGGTGGCTCTTACAT	TCACGT	822			
QY	421	TTATTTCTGGTGGG	CGACGAGACATATCTA	AAAAATTGAAGAAATTTT	TCA	480			
Db	823	TTATTTCTGGTGGG	CGACGAGACATATCTA	AAAAATTGAAGAAATTTT	TCA	882			
QY	481	TGCTATCATGACAG	AGAGATGGCTGGTTGA	AGTGCATGAGCTTGGGAG	CTTAACAC	540			
Db	883	TGCTATCATGACAG	AGAGATGGCTGGTTGA	AGTGCATGAGCTTGGGAG	CTTAACAC	942			
QY	541	CGGCTGCACAGAGT	GTCTCCAAATCATGA	AGCAATTTGGCAGACAAAT	TGGAATGT	600			
Db	943	CGGCTGCACAGAGT	GTCTCCAAATCATGA	AGCAATTTGGCAGACAAAT	TGGAATGT	1002			
QY	601	CTTTCACTCAATAG	CACATTTTACCGGTT	TATATAGGGGTTTACACG	TGGTGA	660			
Db	1003	CTTTCACTCAATAG	CACATTTTACCGGTT	TATATAGGGGTTTACACG	TGGTGA	1062			
QY	661	GCTAACCCCTGGAT	TTTGGCATCACCCCTG	TTCTTGACCTTTAGCCG	CCCATCTGGCC	720			
Db	1063	GCTAACCCCTGGAT	TTTGGCATCACCCCTG	TTCTTGACCTTTAGCCG	CCCATCTGGCC	1122			
QY	721	AAAGATCTATCTCAT	TACTATTAAGAAACCT	GTGGCTTGCAGAAAGCTG	GAGCT	780			
Db	1123	AAAGATCTATCTCAT	TACTATTAAGAAACCT	GTGGCTTGCAGAAAGCTG	GAGCT	1182			
QY	781	AGCTGAAAGATCTT	TACACGATCAAGACTG	ATTTGCTTTGGAGACAA	AGAAAGA	840			
Db	1183	AGCTGAAAGATCTT	TACACGATCAAGACTG	ATTTGCTTTGGAGACAA	AGAAAGA	1242			
QY	841	ACTTGAAGGTACACA	AAAAATTTAGAAAGAC	TCTAAAGAAATTTGGGAT	TTAAAGAAAGTAT	900			
Db	1243	ACTTGAAGGTACACA	AAAAATTTAGAAAGAC	TCTAAAGAAATTTGGGAT	TTAAAGAAAGTAT	1302			
QY	901	CACGGCAAAATTTCA	TATGGTGGCGCTTTCT	TATGATCTATCATCATATG	CTCTGGC	960			
Db	1303	TACAGCAAAATTTCA	TATGGTGGCGCTTTCT	TATGATCTATCATCATATG	CTCTGGC	1362			

QY	961	TTTCTGGGATGGGACCTCTTGCTCTCCATGTAATTTCTATTGGACAACTACTCAC	1020
Db	1363	CTTCGGTATGGGACCACTTGCTCTCCATGAGGGAATTTCTATTGGACAACTACTCAC	1422
QY	1021	TGCTCTCTTTCTGTATTAATTTGGGGCTTTAGATTAATGGACAGGCATCCCCAAGCATTTGA	1080
Db	1423	TGTATTTCTTTCTGTATTAATTTGGGGCTTTAGAGTTGGACAGGCATCTCCAAAGCATTTGA	1482
QY	1081	AGCATTTTCAAAAGCAGAGAGAGCAGCTTATGAATCTTCACATTAATTGCACATTAACC	1140
Db	1483	AGCATTTTCAAAAGCAGAGAGAGCAGCTTATGAATCTTCACAGATTAATTGAATTAAGCC	1542
QY	1141	AAGCATTTGACAGCTTTTGGAGAGTGGGCATTAACACAGATTAATTGAAGGGAATTTTGA	1200
Db	1543	AAGTATTTGACAGCTTTTGGAGAGTGGGCATTAACACAGATTAATTGAAGGGAATTTTGA	1602
QY	1201	ATTTCAAAAATGTTCACTTCAGTTTACCCTTCTCGAAAAGAGTTAAATCTTTAAAGGCTCT	1260
Db	1603	ATTTCAAAAATGTTCACTTCAGTTTACCCTTCCTCGAAAAGAGTTAAATCTTTAAAGGCTCT	1662
QY	1261	CAACTGAAAGGTTTCAGAGTGGGCACACAGTGGCGCTGTGGGAAACAGTGGCTCGGGAA	1320
Db	1663	GAACCTGAAAGGTTTCAGAGTGGGCACACAGTGGCGCTGTGGGAAACAGTGGCTCGGGAA	1722
QY	1321	GAGCAGACCGCTGACAGCTATGCAGAGGCTATAGACCCACAGTGGCATGCTGTAT	1380
Db	1723	GAGCAGACCGCTGACAGCTATGCAGAGGCTATAGACCCACAGTGGGGAATGCTGTAT	1782
QY	1381	TGATGGACAGGACATTAAGACCAATTAATTAAGGACCTTCGCGGAAATTTACTGCTGTGT	1440
Db	1783	TGATGGACAGGATTAAGACCAATTAATTAAGGACCTTCAGCGGAAATCATTTGCTGTGT	1842
QY	1441	GAGTCAGAGACCTGTGTGTTTTGCCACACAGATAGCTGAAAACATTCGTAATGGCCGGGA	1500
Db	1843	GAGTCAGAGAACCTGTATGTTTTGCCACACAGATAGCTGAAAACATTCGTAATGGCCGGGA	1902
QY	1501	AAATGTCAACATGGATGAGATGAGAAAGCTGTTAAGGAAGCCAAATGCTATGATTTAT	1560
Db	1903	AAATGTCAACATGGATGAGATGAGAAAGCTGTTAAGGAAGCCAAATGCTATGATTTAT	1962
QY	1561	CATGAACCTACCTAATTAATTTGCACTCTGTTTGGAGAGAGAGGGGCCACAGTGAATGG	1620
Db	1963	CATGAACCTACCTAATTAATTTGCACTCTGTTTGGAGAGAGAGGGGCCACAGTGAATGG	2022
QY	1621	TGACAGAGAAAGAGATCGCATTTGTCGGGGCCCTGTTGCCAGACCCCAAGATTTCTCT	1680
Db	2023	TGGGAGAGAGAGAGGATCGCATTTGTCAGCTGCCCTGTGTCGAACCCCAAGATTTCTCT	2082
QY	1681	GCTGATAGGACCAAGTACGCTTGCACACTGAAAGTGAAGACAGTGTAGGTGGCCCT	1740
Db	2083	GCTGATAGGACCAAGTACGCTTGCACACAGAAAGCAAGACAGTGTAGGTGGCCCT	2142
QY	1741	GGATTAAGCCAGAAAAGCCCGGACCTACCATTTGTATAGCTCATCGTTTGTCTACAGTTTC	1800
Db	2143	GGATTAAGCCAGAAAAGGTCGGACACCATTTGTATAGCTCATCGTTTGTCTACAGTTTC	2202
QY	1801	TAAATGCCATGTCTATGCTGTGTTTTGATGATGAGTCAATTTGTGAGAAAAGAAATCATGA	1860
Db	2203	TAAATGCCATGTCTATGCTGTGTTTTGATGATGAGTCAATTTGTGAGAAAAGAAATCATGA	2262
QY	1861	TGAACTCATGAAAGAGAGGAGATTTACTCTCAAACTTGTCAATTCGACGACAAAGAGAA	1920
Db	2263	TGAACTCATGAAAGAGAGGAGATTTACTCTCAAACTTGTCAATTCGACGACGAGAGAA	2322
QY	1921	TGAAATTTGATTAAGAAAATGGCACTGTGTAATCCAAAAGTGAAGTAGTGTCTTGAAT	1980
Db	2323	TGAAATTTGATTAAGAAAATGGCACTGTGTAATCCAAAAGTGAAGTAGTGTCTTGAAT	2382
QY	1981	GTCTCCAAAAGATTCAGGCTCAGTTTAATTAATAAGAAATCACTCCAGAGAGATTTCA	2040
Db	2383	GTCTCCAAAATGATTTCAAGATTCAGTCTAATTAAGAAAAGATCACTCCAGAGAGTCTCG	2442
QY	2041	TGCACCAAGGCAAGACAGAAAGCTTGTCAAAAAGAGCACTGAAATGAGAAATGTACC	2100

Db	2443	TGATTCACAAACCCCAACAGAAAGAGCTTACTACCAAGAGGCTCTGGATGCAAGATATACC	2502
OY	2101	TCGAGTTTCCCTCTGGAGGATTTGGAAGCTGAACTCAACTGAATGGCCCTTATTTGTGGT	2160
Db	2503	TCGACTTTCCTTTTGGAGGATTTATGAAAGCTAAATTTAACTGAATGGCCCTTATTTGTGG	2562
OY	2161	TGATATATTTTGGCTATTTATTAACGAGAGCCCTGCACACAGCAATTTTCATATATTTTC	2220
Db	2563	TGGTATTTTGGTGCATTTATTAATGAGAGGCTGCACACAGCAATTTTCATATATTTTC	2622
OY	2221	AAGGATTTATGAGGATCTTACCCGAGATGAGATCTCGAAACAACAAAGACAGATAGTAA	2280
Db	2623	AAAGATTTATGAGGATTTTATACAGAAATGATAGCTCGAAACAACAAAGACAGATAGTAA	2682
OY	2281	CATGTTTCTGATTTGTTTCTAGTCTTGGAATTTTCTTTTTCATATTTTCCCA	2340
Db	2683	CTTGTTTTCAATTTGTTTCTAGCCCTTGGAATTAATTTCTTTTATTTTCTTTCA	2742
OY	2341	GGGCTTCACATTTGGCAAAAGCTGGGAGATCCTCAATAGGCGCTTCGATACATGTTTT	2400
Db	2743	GGGTTTTCACATTTGGCAAAAGCTGGAGATCCTCCACAAAGGCGCTCGATACATGTTTT	2802
OY	2401	CAGATCCATGCTGAGACAGAGATGTCACTGGTTTATGACCTTAACCCATAAACACACTGGAGC	2460
Db	2803	CGATCCATGCTGAGACAGAGATGTGAGTTGTTTATGACCTTAACCAACACACTGGAGC	2862
OY	2461	ATTGACAAACAGGCTTGCCAATGATCCGAGCTCAAGTTAAAGGGGCTATAGTTCCAGGCT	2520
Db	2863	ATTGACTACAGGCTTGCCAATGATCCTCTCAAGTTAAAGGGGCTATAGTTCCAGGCT	2922
OY	2521	TGCTGTCATTTACCAGAAATATATAGCAAAATCTTTGGGACAGGCAATTAATATCCTTAATCTA	2580
Db	2923	TGCTGTAATTTACCAGAAATATATCAAAATCTTTGGGACAGGAAATTAATATCCTTAATCTA	2982
OY	2581	TGGTGGCAATTTACACTTTTACTTCTTAGCAATTTGTACCATCATTTGCAATAGCAGAGT	2640
Db	2983	TGGTGGCAACTACACTGTACTTACTCTTAGCAATTTGTACCATCATTTGCAATAGCAGAGT	3042
OY	2641	TGTTGAATGAAAAATGTTGTCTGGACACAGCTGCAAAAGATTAAGAAAGACTTGAAGAGAC	2700
Db	3043	TGTTGAATGAAAAATGTTGTCTGGACACAGCAAGAAAGATTAAGAAAGACTTGAAGAGTGC	3102
OY	2701	TGGGAAGATTGCTCTGACAAAGCATGCAAACTCCGAACTGTTGTTCTTGTACTCAGGA	2760
Db	3103	TGGGAAGATGCTCTGACAAAGCATAGAAACTTCGAACTGTTGTTCTTGTACTCAGGA	3162
OY	2761	GCAGAAAGTTGAATACATGTAATGACACAGAGTTTGCAGATGCCATACAGAAACCTTTGAG	2820
Db	3163	GCAGAAAGTTGAACATATGTAATGCTCAGAGTTTGCAGATGCCATACAGAAACCTTTGAG	3222
OY	2821	GAAAGCAACACTCTCCGGGGCTCATATTTTCTATCACCAGGCAATGATATATTTTCTTA	2880
Db	3223	GAAAGCAACACTCTTGGAAATTAATATTTTCTTCCACCAAGCAATGATATATTTTCTTA	3282
OY	2881	TGCTGGCTGTTTCCGTTTGGTGGCTACTTGTGGCAATATGATTCATGAAACTTTTCAGA	2940
Db	3283	TGCTGGAGTTTCCGTTTGGAGGCTACTTGTGGCAATATTAACCTATGAGACTTTGAGGA	3342
OY	2941	TGTTCTTTTGGATCTCAGCATTTGTGTTGGTGCATATGGCACTGGGGCAGGTCACTTC	3000
Db	3343	TGTTCTGTTAGTAATTTTCAAGCTGTGTGTTTGTGGTCCATGGCCGTGGGGCAAGTCACTTC	3402
OY	3001	ATTGTCCTCTGACTATGGCAAAAGCCAAAGATATCAGACAGCCCACTCATATGATCATTTGA	3060
Db	3403	ATTGTCCTCTGACTATGGCAAAAGCCAAATATCAGACAGCCCACTCATATGATCATTTGA	3462
OY	3061	AAAAAGCCCTGTGATTTGACAGTACAGCCCTCAGGCGCTCAAGCAACAAATATGTTTGAAG	3120
Db	3463	AAAAAGCCCTGTGATTTGACAGTACAGCAAGGAGGCTTAATGCCAAACATTTGGAAG	3522
OY	3121	AAATGCACTTTAATGAGAGTGGTTCACATATCCATCGACCAAGCATCCCGTCT	3180

Db 3523 AAATGTCACATTTGGTGAAGTTGTATTCACATTCCTCCACCCGACGACATCCCAAGTCT 3582
Qy 3181 CCAGGGGCTGAGGCTCGAGGTGTAAGAGGGCCAGACGCTGGCCCTCTAGGTAGCACTGG 3240
Db 3583 TCAGGAGCTGAGGCTGGAGGTGAAGAGGGCCAGACGCTGGCTCTGGTGGGACAGCATGG 3642
Qy 3241 CTGTGGGAAGACACAGTGTTCAGCTCTGAGGCTGAGGCTCTGATGACCCCTTGGCTGTT 3300
Db 3643 CTGTGGGAAGACACAGTGTTCAGCTCTGAGGCTGAGGCTCTGATGACCCCTTGGGAGAA 3702
Qy 3301 AGTGTATGTATGGCAAGAGATTAAGACACCTGTAATGTCCAGTGGTCCGACACACCT 3360
Db 3703 AGTGTGTGTATGGCAAGAGATTAAGACACCTGTAATGTCCAGTGGTCCGACACACCT 3762
Qy 3361 GGGGATCTGTCTGAGAGGACCATCTGTTGACTGTGAGATTCGCGAGAACCTTGCCTA 3420
Db 3763 GGGGATCTGTCTGAGAGGACCATCTGTTGACTGTGAGATTCGCGAGAACCTTGCCTA 3822
Qy 3421 TGGAGACACAGCCGGGCTGTATTCACATGTAAGATTAATGAGAGGACGACAGAGGCCAA 3480
Db 3823 TGGAGACACAGCCGGGCTGTATTCACATGTAAGATTCGAGAGGACGACAGAGGCCAA 3882
Qy 3481 CATACACACTTCATGAGACACTCCCTGTAATTAACAACACAGTGAAGACAAAG 3540
Db 3883 CATACATGCTTCATGAGACACTCCCTGTAATTAATTAATGACTAAAGTGAAGACAAAG 3942
Qy 3541 AACCCAGCTCTGTGTGGCCAGAAAGCCGATTCGATAGCTCGCGCTGTGAGACA 3600
Db 3943 AACCCAGCTCTGTGTGGCCAGAAAGCCGATTCGATAGCTCGCGCTGTGAGACA 4002
Qy 3601 GCCATATTTTCTTTTGGATGAAGCTACATCAGCTCGATACAGAAAGTGAAGAGT 3660
Db 4003 GCCATATTTTCTTTTGGATGAAGCTACATCAGCTCGATACAGAAAGTGAAGAGT 4062
Qy 3661 TGTCCAAAGAGCCCTGGAGCAAAAGCCAGAGAGCCGACCTGATTTGATGATGCCGCCG 3720
Db 4063 TGTCCAAAGAGCCCTGGAGCAAAAGCCAGAGAGCCGACCTGATTTGATGATGCCGCCG 4122
Qy 3721 CTGTTCACATTCAGAAATGAGATTAATGATGATTTGATGATGAGAGCAAGTCAAGGA 3780
Db 4123 CTGTTCACATTCAGAAATGAGATTAATGATGATTTGATGATGAGAGCAAGTCAAGGA 4182
Qy 3781 GGATGGACACATCAACAGCTGTGGCCAGAAAGGATCTAATTTTCCATGGCAGTGT 3840
Db 4183 GGATGGACACATCAACAGCTGTGGCCAGAAAGGATCTAATTTTCCATGGCAGTGT 3900
Qy 3841 CCAGGCTGAGCAAAAGCCGTGAGTGTGAGTGTGAGTGTGAGTGTGAGTGTGAGTGT 4242
Db 4243 CCAGGCTGAGCAAAAGCCGTGAGTGTGAGTGTGAGTGTGAGTGTGAGTGTGAGTGT 3900
Qy 3901 ATTGTGTAAACATGAGATTAATCAAGATTAAGATTAAGATTAAGATTAAGATTAAG 4302
Db 4303 ATT--TGTTTGATATGACATTTATTCAGTTAA--ACCAACACTTTCAGCAAAAT 3960
Qy 3961 ATTGTAAAGCTGTTTAACTTCTGCTGCACTGAGATCATTCACCAAGTTGAC 4355
Db 4356 ATGAAGAGATCTGTTTAACTTCTGCTGCACTGAGATCATTCACCAAGTTGAC 4402
Qy 4021 AGCTTCAAGATTTATTAATTAAGAACCA--AAGAAATATTCATGAGATTAAG 4403
Db 4404 ----AGAGACTTCGTATTAAGAACCAAGAGATGAGACATCATCACTGAGAGAAAT 4078
Qy 4079 ACTGTGTAAATGCAATTAATAAATTAATAGATTAATCAAGTATTTGTTAATAAT 4458
Db 4459 CATAGTTTAACTGATTAATTAATTAATTAAGCAATTAAGTATTAATTAAGATTA 4138
Qy 4139 TGTATATTTTGTATATATTT-----ATTGTACTATGCTGCTGCTGAAGAT 4518
Db 4519 ATGTGTAATTTTGTATATATTTTCCCATTTGAGACTGTAACAGCTGCTCTAAAGAT 4190
Qy 4191 TATAGAGTGTGTAAGAGACTG--AATGTTGATTAAGAGTGTGATTAATAAGCTAA 4578
Db 4579 TATAGAGTGTGTAAGAGATTAAGATTTGATTAAGATG--TCTATATTAATAAGCTAA 4655

Qy 4250 ACTTTTATAT 4259
Db 4636 ACTTTCAATGT 4645
RESULT 14
AX336420
LOCUS AX336420
DEFINITION Sequence 6929 from Patent WO0194629.
ACCESION AX336420
VERSION AX336420.1 GI:18127139
KEYWORDS
SOURCE human.
ORGANISM Homo sapiens
REFERENCE Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
1
AUTHORS Young, P. E., Augustus, M., Carter, K. C., Edner, R., Endress, G.,
Horrikan, S., Soppet, D. R. and Weaver, L.
TITLE Cancer gene determination and therapeutic screening using signature
JOURNAL Patent: WO 0194629-A 6929 13-DEC-2001;
FEATURES
source location/Qualifiers
1..4646
BASE COUNT 1371 a 892 c 1129 g 1254 t
ORIGIN
Query Match 77.7%; Score 3324.4; DB 6; Length 4646.
Best Local Similarity 87.6%; Pred. No. 0;
Matches 3740; Conservative 0; Mismatches 486; Indels 44; Gaps 8;
Qy 1 GGAGCGCGAGTGGGATGATGATCCGGAAGAGCGCGTGAAGGAGTCAAGAGAACTT 60
Db 409 GGAGCGCGAGTGGGATGATGATCCGGAAGAGCGCGTGAAGGAGTCAAGAGAACTT 60
Qy 61 CTGGAATAATGGGCAAAAAAATGAATAAAGAGAGAGAGAGAGAGAGAGAGAGAGAG 120
Db 469 CTGGAATAATGGGCAAAAAAATGAATAAAGAGAGAGAGAGAGAGAGAGAGAGAGAG 120
Qy 121 CAGCTTCATGATTTGCTGCTATTCATTAATGCTGATGATGATGATGATGATGATGAT 180
Db 529 TGTATTTTCAATGTTGCTGATTCATTAATGCTGATGATGATGATGATGATGATGAT 588
Qy 181 AATGGCTGCATTCATTCATGAGTGCAGTCCCTCATGATGATGATGATGATGATGATGAT 240
Db 589 TTTGGCTGCATTCATTCATGAGTGCAGTCCCTCATGATGATGATGATGATGATGATGAT 648
Qy 241 GACGATAGCTTTGCAATGACAGATTTTCAAGAAACAACTTTCAGTTATATTA 300
Db 649 GACGATAGCTTTGCAATGACAGATTTTCAAGAAACAACTTTCAGTTATATTA 300
Qy 301 TGAAGATTAATGACAAATACACATTTTCATCAACCATTCGAGAGAGAGAGAGAGAG 360
Db 703 TGAAGATTAATGACAAATACACATTTTCATCAACCATTCGAGAGAGAGAGAGAGAG 762
Qy 361 GTATGCTATTAATTAAGTGGATCGTGTGCTGCTGATGATGATGATGATGATGATGATGAT 420
Db 763 GTATGCTATTAATTAAGTGGATCGTGTGCTGCTGATGATGATGATGATGATGATGATGAT 822
Qy 421 TTGATTCGCTGCTGCGAGAGAGAGAGAGAGATTCATCAAAATTAATAAACAATTTTCA 480
Db 823 TTGATTCGCTGCTGCGAGAGAGAGAGAGATTCATCAAAATTAATAAACAATTTTCA 882
Qy 481 TGTATATGAGAGAGAGATGCTGTTGATGATGATGATGATGATGATGATGATGATGATGAT 540
Db 883 TGTATATGAGAGAGAGATGCTGTTGATGATGATGATGATGATGATGATGATGATGATGAT 942
Qy 541 CCGGCTCAGAGAGATGCTCCAAATCAATGAAGATTTGGCAAAATTTGAATGTT 600

Db 943 CCGACCTACAGATGATGCTCTCTAAGATTATGAAGTATTGGTGCACAAAATTGGAATGTT 1002
Qy 601 CTTTCACTAATAAGACAATTTTACCGGTTTATAGTGGGTTTACACGTGGTGGAA 660
Db 1003 CTTTCACTAATAAGACAATTTTACCGGTTTATAGTGGGTTTACACGTGGTGGAA 1062
Qy 661 GCTAACCTTGTGATTTTGGCCATCAGCCCTGTCTTGGACTTTTACGCCGCTGTGGGC 720
Db 1063 GCTAACCTTGTGATTTTGGCCATCAGCTCTTCTTGGACTTTTACGCCGCTGTGGGC 1122
Qy 721 AAGATATCTCTCTCAATTACTGATTAAGAACTTTGGCTATGCAAAAAGCTGAGCAGT 780
Db 1123 AAGATATCTCTCTCAATTACTGATTAAGAACTTTAGGCTATGCAAAAAGCTGAGCAGT 1182
Qy 781 AGCTGAAGAGTCTTGAACCAATCAAACTGTGATGCTTGGAGAGCAAAAAGAA 840
Db 1183 AGCTGAAGAGTCTTGAACCAATCAAACTGTGATGCTTGGAGAGCAAAAAGAA 1242
Qy 841 ACTTGAAGAGTCTTGAACCAATTAAGAAAGCTTAAGAAATTTGGATTAAGAAAGCTAT 900
Db 1243 ACTTGAAGAGTCTTGAACCAATTAAGAAAGCTTAAGAAATTTGGATTAAGAAAGCTAT 1302
Qy 901 CACGCCCAACATTTCTATTTGTCGCCCTTCTTATTTGATCTATGCTATGCTCTGCG 960
Db 1303 TACAGCCAAATTTCTATTTGATGCTCTTCTGCTGATCTATGCTATGCTCTGCG 1362
Qy 961 TTTCTGATGTGGACCTTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1020
Db 1363 CTTCTGATGTGGACCTTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1422
Qy 1021 TGTCTCTCTTCTGATTAATTTGGGCTTTTATTTGATTTGACAGGACTCCCAAGCAATGA 1080
Db 1423 TGTATCTCTTCTGATTAATTTGGGCTTTTATTTGATTTGACAGGACTCCCAAGCAATGA 1482
Qy 1081 AGCATTTGCAAAACGAAGAGAGAGAGCTTATGAATCTTCAAGTATTTATCAATAAAC 1140
Db 1483 AGCATTTGCAAAACGAAGAGAGAGAGCTTATGAATCTTCAAGTATTTATTAAGAGG 1542
Qy 1141 AAGCATTTGCAAGCTATTTGAAGAGTGAACATTAACACAGATTAATTTAAGGAAATTTGGA 1200
Db 1543 AAGTATTTGCAAGCTATTTGAAGAGTGAACATTAACACAGATTAATTTAAGGAAATTTGGA 1602
Qy 1201 ATTCAAAATTTGCTACCTTATGATTAACCTTCTGAAAGAGTAAAGTCTTAAAGGCTCT 1260
Db 1603 ATTCAAAATTTGCTACCTTATGATTAACCTTCTGAAAGAGTAAAGTCTTAAAGGCTCT 1662
Qy 1261 CAACCTGAAGGTTCAAGATGGGACAGACAGTGGGCTGTTGGGAAACATGCTGCGGAA 1320
Db 1663 GAACCTGAAGGTTCAAGATGGGACAGACAGTGGGCTGTTGGGAAACATGCTGCGGAA 1722
Qy 1321 GAGACAGACCTGCTGATGTGACAGAGGCTCTATGACCCACAGATGGCATGCTGTAT 1380
Db 1723 GAGACAGACCTGCTGATGTGACAGAGGCTCTATGACCCACAGATGGCATGCTGTAT 1782
Qy 1381 TGAATGAGAGAGATTTGAGCAATTAATGTAAGGCAATCTTGGGAAATTAATGCTGTGT 1440
Db 1783 TGAATGAGAGAGATTTGAGCAATTAATGTAAGGCAATCTTACGGGAAATTAATGCTGTGT 1842
Qy 1441 GAGTCAGAGAGCTGTGTTGTTGGCCACACAGATGAGTGAAGAAATTCGTTATGCGCGGA 1500
Db 1843 GAGTCAGAGAGCTGTGTTGTTGGCCACACAGATGAGTGAAGAAATTCGTTATGCGCGGA 1902
Qy 1501 AAATGTCAACATTTGATGATGATGGAAGCTGTTAAGGAAGCCATGCTATGATTTAT 1560
Db 1903 AAATGTCAACATTTGATGATGGAAGCTGTTAAGGAAGCCATGCTATGATTTAT 1962
Qy 1561 CATGAACCTACCTAATAAATTTGACACTCTGTGTTGGAGAGAGAGGGGCCAGTGTAGTGG 1620
Db 1963 CATGAACCTACCTAATAAATTTGACACTCTGTGTTGGAGAGAGAGGGGCCAGTGTAGTGG 2022
Qy 1621 TGGACAGAAAGAGAACTGCGCATTTGCTGGGCTGCTGTTGCGAACCCCAAGATTTCTCT 1680
Db 2023 TGGGAGAAAGAGAGAACTGCGCATTTGCGCATTTGCGGCTGCTGCGAACCCCAAGATTTCTCT 2082

Qy 1681 GCTGATGAGGCACGTCAGCTGTGACACTGAAGTGAAGCAAGTGGTTCAGGTGGCCCT 1740
Db 2083 GCTGATGAGGCACGTCAGCTGTGACACTGAAGTGAAGCAAGTGGTTCAGGTGGCCCT 2142
Qy 1741 GATTAAGGCACAAAAGCCGAGTACATTTGATAGCTATCTGTTGCTCTACATGTCG 1800
Db 2143 GATTAAGGCACAAAAGCCGAGTACATTTGATAGCTATCTGTTGCTCTACATGTCG 2202
Qy 1801 TAAATGCGATGTCATTTGCTGTTTGTATGATGAGAGTCAATTTGGAGAAAGAAATCATGA 1860
Db 2203 TAAATGCGATGTCATTTGCTGTTTGTATGATGAGAGTCAATTTGGAGAAAGAAATCATGA 2262
Qy 1861 TGAACATGAAGAAAGAGAGGCAATTTACTTCAAACTTGTCAATGTCAGACAAAGGAAA 1920
Db 2263 TGAACATGAAGAAAGAGAGGCAATTTACTTCAAACTTGTCAATGTCAGACAAAGGAAA 2322
Qy 1921 TGAATTTGATTAAGAAATGCCACCTGTGTAATCCAAAAGTGAAGTGAAGCTTGGAAAT 1980
Db 2323 TGAATTTGATTAAGAAATGCCACCTGTGTAATCCAAAAGTGAAGTGAAGCTTGGAAAT 2382
Qy 1981 GTCTCCAAAGATTCAGGGGTCAGTTTAAATTAAGAAAGATCACTCGCAGAGATATCA 2040
Db 2383 GTCTCCAAAGATTCAGGGGTCAGTTTAAATTAAGAAAGATCACTCGCAGAGATATCA 2442
Qy 2041 TGCACCAAGGCGCACAGAGAAAGCTTGTGTACAAAAGAGACTGGAATGAGATGTAC 2100
Db 2443 TGCATCAAGGCGCACAGAGAAAGCTTGTGTACAAAAGAGACTGGAATGAGATGTAC 2502
Qy 2101 TCCAGTTTCTCTTGGAGAGATTCGAGAGTGAAGTCACTCACTCACTGATTTGTTGT 2160
Db 2503 TCCAGTTTCTCTTGGAGAGATTCGAGAGTGAAGTCACTCACTCACTGATTTGTTGT 2562
Qy 2161 TGTATTTTGTGCTATTTAATTAAGAGAGGCTGCAACACGATTTTCAATATATTTTC 2220
Db 2563 TGTATTTTGTGCTATTTAATTAAGAGAGGCTGCAACACGATTTTCAATATATTTTC 2622
Qy 2221 AAGGATTAAGGATCTTATCCGAGATGAGATCTGAAACAAAGAGAGAGATTAAGTA 2280
Db 2623 AAGGATTAAGGATCTTATCCGAGATGAGATCTGAAACAAAGAGAGAGATTAAGTA 2682
Qy 2281 CATGTTTCTGATTTGTTCTGATGCTTGAATTTTCTTTATTTTCAATTTTCCCA 2340
Db 2683 CATGTTTCTGATTTGTTCTGATGCTTGAATTTTCTTTATTTTCAATTTTCCCA 2742
Qy 2341 GGGCTTCAATTTGGCAAAAGCTGGGAGATCTCACTAAGCGGCTTGATACATGTTT 2400
Db 2743 GGGCTTCAATTTGGCAAAAGCTGGGAGATCTCACTAAGCGGCTTGATACATGTTT 2802
Qy 2401 CAGATCCATGCTGAGACAGAGATGTACGCTGTTGATGACCTTAAACACCACTGGAGC 2460
Db 2803 CAGATCCATGCTGAGACAGAGATGTAGTGGTTGATGACCTTAAACACCACTGGAGC 2862
Qy 2461 ATTGACACAGGCTTGGCAATGATGGGCTCAAGTTAAAGGGGCTATAGGTTCCAGGCT 2520
Db 2863 ATTGACACAGGCTTGGCAATGATGGGCTCAAGTTAAAGGGGCTATAGGTTCCAGGCT 2922
Qy 2521 TGGCTATTTACCCAGAAATATAGCAAACTTTGGAGACAGGCAATTTATATATCTTAATCTA 2580
Db 2923 TGGCTATTTACCCAGAAATATAGCAAACTTTGGAGACAGGCAATTTATATATCTTAATCTA 2982
Qy 2581 TGGTTGCAATTAACATTTTACTCTTACCAATTTGTAATGACCAATTTGCAATGACAGAGT 2640
Db 2983 TGGTTGCAATTAACATTTTACTCTTACCAATTTGTAATGACCAATTTGCAATGACAGAGT 3042
Qy 2641 TGTGGAATGAAGAAATGTTGCTGTGACAAAGCACTGAAGATTAAGAAAGAGCTGAAGAGG 2700
Db 3043 TGTGGAATGAAGAAATGTTGCTGTGACAAAGCACTGAAGATTAAGAAAGAGCTGAAGAGG 3102
Qy 2701 TGGGAAGATTTCTACAGAGAGCTGCAAACTTCCGAAGCTGTTCTTTGACTCGGGA 2760
Db 3103 TGGGAAGATTTCTACTGTGAAGCAATGAAGAACTTCCGAAGCTGTTCTTTGACTCGGGA 3162

QY	2761	CGAGAGATTGAATACATGATGACACAGAGATTGCAAGTACCATACAGAACTCTTTAG	2820
Db	3163	GCAGAAAGTTGGAACATATATGATGCTCAGAGTTTGCAGGATACATACAGAACTCTTTAG	3222
QY	2821	GAAGACACACATCTTGGGGGTCTCATTTTCTATCACCCAGGCATATGATTTTTCCTA	2880
Db	3223	GAAGACACACATCTTTGGATTATCAATTTTCTTCCATCCAGGCAATGATGATTTTTCCTA	3282
QY	2881	TGCTGGCTGTTTCCGGTTTGGTGCCACTTGGTGGCAATATGTTTCATGAACTTTTCAGA	2940
Db	3283	TGCTGGATGTTTCCGGTTTGGAGCCACTGGTGGCACATTAACCTATGAGACTTTTGGGA	3342
QY	2941	TGTTCTTTTGGATTCTCAGCATTTGTCTTTGGTCCATATGGCAGTGGGGCAGGTCACTTC	3000
Db	3343	TGTTCTGTTATGATTTTTCAGCTGTTGTCTTTGGTGCCATATGGCCCTGGGGCAGTCACTTC	3402
QY	3001	ATTTGCTCTGACATATGCCAAGGCCAAGATATCAGACACCACAGTCAATGATCATTTGA	3060
Db	3403	ATTTGCTCTGACATATGCCAAGGCCAAMATATCAGACACCACATCATATCATATTTGA	3462
QY	3061	AAAAACCCCTCGATTGACAGTACAGCCCTCACAGCCCTCAAGCCAAATACGTTGGAAAG	3120
Db	3463	AAAAACCCCTTGATTGACAGCTACAGCTACAGAGGCCAAGGCTATATCCCAACATTTGAAG	3522
QY	3121	AAATGTGACATTTAATGAGGTCTGTTCAATATCCACATCTGACACACAGACATCCCGTCT	3180
Db	3523	AAATGTCACATTTGGTGAAGTTGTAATTCACATATCCACCCGACCGGACATCCCACTGT	3582
QY	3181	CCAGGGGCGTAGCCCTCGAGGTGAAGAGGCGCACAGCTGGCCCTGTAGTACAGTGG	3240
Db	3583	TCAGGAGATGAGCTCGAGGTGAAGAGAGGCGCACAGCTGGGCTGTGTGGCAGCAGTGG	3642
QY	3241	CTGTGGGAAGACACAGTTGTTACGCTCTAGAGGCCCTCTATGACCCCTTGGCTGGTTC	3300
Db	3643	CTGTGGGAAGACACAGTGTGTCAGCTCTAGAGGCCCTCTATGACCCCTTGGCTGGTTC	3702
QY	3301	AGTGTAAATTATGGCAAAAGATTAAGACACTGAATGTCCAGTGGCTCCGACACACT	3360
Db	3703	AGTGTGCTTATGACCAAAAGATTAAGAGCAGTATGTTCACTGGCTCCGACACACT	3762
QY	3361	GGGCACTGTGCTCAGAGAGCCCATCCCTGTTGATCGACAGCATATGGCCGAACATTGCCCTA	3420
Db	3763	GGGCACTGTGCTCAGAGAGCCCATCCCTGTTTACTGTCAGCATATGGTGAACATTGCCCTA	3822
QY	3421	TGGAGACAACAGCCGGGTGCTATCACATGAGATTAATGACAGGACGACAGAGGCCAA	3480
Db	3823	TGGAGACAACAGCCGGGTGCTCTACAGGAAAGATCTGTAGGGCAGAGAAAGAGGCCAA	3882
QY	3481	CATACACCACTTTCATGAGACACTCCCTGAGAAATATGACAACCCAGATAGAGACAAAG	3540
Db	3883	CATACACTGCTTCATGAGACTACTGCTCTAATATATGACATAAATAGAGACAAAG	3942
QY	3541	AATCCAGCTCTCGGTGGGCGCAAAACAGCGCATTTGCCATAGCTCGCCTTGTATTGACA	3600
Db	3943	AATCTACGCTCTCTGGGGCGCAAAACAGCGATTTGCCATAGCTCGCCTTGTATTGACA	4002
QY	3601	GCCCTATATTTGCTTTTGGATGAAAGTACATATCAGCTCTGGATATGAGAAAGTAAAAAGT	3660
Db	4003	GCCCTATATTTGCTTTTGGATGAAAGCACACTCAGCTCTGGATATGAGAAAGTAAAAAGT	4062
QY	3661	TGTCCAAAGAGCCCTGGGCAAAAGCAGAGAGAGGCCGCACTGCATTTGTATGCCACCG	3720
Db	4063	TGTCCAAAGAGCCCTGGGCAAAAGCAGAGAGAGGCCGCACTGCATTTGTATGCCACCG	4122
QY	3721	CTTGTCCACATTCAGATGAGATTTAATATAGGTGTTTCAGATATGCGAAAGTCAAGGA	3780
Db	4123	CTTGTCCACATTCAGATGAGACTTAATATAGGTGTTTCAGATATGCGAAAGTCAAGGA	4182
QY	3781	GCATGGACACATATCAACAGCTGCTGGGCCAAGAGGCATATTTTTCATATGCTCAGAGT	3840
Db	4183	GCATGGACACATATCAACAGCTGCTGGCACAGAAAGGCATATTTTTCATATGCTCAGAGT	4242
QY	3841	CCAGGCTGGAGCAAGGCGTAGTGAACCTGTGGCGCATATGAGCTGTTAAATATTTTAAAT	3900

QY	Db	4243	CCAGGCTGGAAACAAAGCGCCAGTGAAC	TCTGACGTATGAGATGTTAAATAC	TTTAAAT	4302
QY	3901	ATTGTGTGTTAAACATGGCATTTATATCAAGT	TTAAAGTGAGGAC	CTTACTG	AAAAAACT	3960
Db	4303	ATT--TGTTTGATATGACATTTATTC	AAAGTTAA	-----AGCAAC	CACTTACGAAAT	4355
QY	3961	ATGTAGAACTAACCGTTTAACTTTCTTCTG	CACTGAAGATTCAT	TTCCACC	AAAGTTGAG	4020
Db	4356	ATGAAGAGGATCTGTTTAACTTTCTG	ATCAAGTTCAGAGTCTC	-----		4403
QY	4021	AGTCTTCAGATTTTATATTAATAAGAACCA	CCAA--AAGAA	CACTTATCTGAT	TGGAATAAAT	4078
Db	4404	----AAGAGACTTCGTATTAATAAGAACAG	AGAGTGAAGACATCATCAAGTGAAGAAAT			4458
QY	4079	ACTGAGTGTATATGCACTTATTAATAATTA	TATAGAGTAAATTC	CAAGTACATTTTGT	TATATAAT	4138
Db	4459	CATAGTTTAACTGCACTATTAATAATTTAT	TAACGATTAAGATGATTTTAAAGATAA			4518
QY	4139	TGTATATTTTGTGTTATATTTT-----	ATTGTACTTACGCTTCCTG	CAAAAGAT		4190
Db	4519	ANGTGTATTTTGTGTTATATTTTCCATTTG	AGCTGATACGATCTGCTTGCTATAAAGAT			4578
QY	4191	TATAGAAAGTGAATAAAGACTG-AATGTTG	ATATAAAGTGTAGCTATATATAAACTAA			4249
Db	4579	TATAGAAAGTGAATAAAGATTAATGAATG	TTTGCACTAAAGTG---TGTATATATAA	ACTTA		4635
QY	4250	ACTTTTATAT 4259				
Db	4636	ACTTTCATGT 4645				
RESULT 15						
LOCUS AX336708 4646 bp DNA linear PAT 09-JAN-2002						
DEFINITION Sequence 7217 from Patent WO0194629.						
ACCESSION AX336708						
VERSION AX336708.1 GI:18127427						
KEYWORDS human.						
SOURCE Homo sapiens						
ORGANISM Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.						
REFERENCE 1 Young, P. E., Augustus, M., Carter, K. C., Eder, R., Endress, G., Horrigan, S., Soppet, D. R. and Weaver, Z. Cancer gene determination and therapeutic screening using signature gene sets Patent: WO 0194629-A 7217 13-DEC-2001;						
JOURNAL Avalon Pharmaceuticals (US) Location/Qualifiers						
FEATURES						
source 1..4646						
BASE COUNT 1371 a 892 c 1129 g 1254 t						
ORIGIN						
Query Match Best Local Similarity 87.6%; Pred. No. 0; Matches 3740; Conservative 0; Mismatches 486; Indels 44; Gaps 8;						
QY	1	GGAGCGCAGATGCGGATGATGATCCTGAAGGAGGCGCTTAAGGGAGTCCACAGAAACTT	60			
Db	409	GGAGCGCGAGTGGCGGATGATGATCTTGAAGGGGACCCGCAATGGAGACCAAGAGAA	68			
QY	61	CTGGAAATTTGGCCAAAAAAGTAAAAAAATGAGAGAAAGAAAGAAACCACTGTGAG	120			
Db	469	CTTTTAACTGAACCAATTAATAAGTGAAGAAAGATGAAGAAAGAAAGAAACCACTGTGAG	528			
QY	121	CACGTTTGCAATGTTTGGCTATTCAAATTTGGCTGATAGTGTATATATGTGTGGGAGC	180			
Db	529	TGTATTTTCATGTTTTCGATATTCAAATTTGCTTGGACAAAGTGTATATGTGTGGGAGC	588			

181 AATGGCTGCCATCATCTGAGAGCTGCACCTCCCTCATGATGCTGTTTGGAAACAT 240
|||||
589 TTGGCTGCATCATCTGAGAGCTGCACCTCCCTCATGATGCTGTTTGGAAACAT 648
241 GACACATGCTTGGCAATGAGAAATTTCAAGAAACAACTTTCCAGTTATATTA 300
|||||
649 GACAGATATCTTGGCAATGAGAAATTTGAGATCTGATGTCA-----AACATCAC 702
301 TGAAGTATTTACGAACATATACAAACATTTGATACACCTGGAGAGGAAATGACAC 360
|||||
703 TAATAGAGTATATCAATGATACAGGGTCTTCAATGATCTGAGAGAAACATGACAG 762
361 GTATGCTATTTATACAGTGGAGTGGCTGGCTGGCTGGCTGGCTGGCTGGCTGGCT 822
763 GTATGCTATTTATACAGTGGAGTGGCTGGCTGGCTGGCTGGCTGGCTGGCTGGCT 882
421 TTGATCTGCTGCTGGCTGGCTGGCTGGCTGGCTGGCTGGCTGGCTGGCTGGCTGG 882
823 TTGATCTGCTGCTGGCTGGCTGGCTGGCTGGCTGGCTGGCTGGCTGGCTGGCTGG 882
481 TGTATGATGCGACAGAGATGGCTGGCTGGCTGGCTGGCTGGCTGGCTGGCTGGCT 942
883 TGTATATATGCGACAGAGATGGCTGGCTGGCTGGCTGGCTGGCTGGCTGGCTGGCT 942
541 CCGGCTCAGACAGATGCTCTCCAAATCAATGAGAAATTTGGCAAAATTTGAAATGTT 600
|||||
943 CCGACTTACAGATGATGCTCTTAAGATTAAGAAATTTGGTGAACAAATTTGAAATGTT 1002
601 CTTTACATCATATGACATTTTTCACCGGTTTATATGAGGCTTTACACGCTGGTGA 660
1003 CTTTACATCATATGACATTTTTCACCGGTTTATATGAGGCTTTACACGCTGGTGA 1062
661 GCTAACCTTGTGATTTTGGCCATCACCCCTGTTTGGACCTTTCAGCCGCTTGGGC 720
1063 GCTAACCTTGTGATTTTGGCCATCACCCCTGTTTGGACCTTTCAGCCGCTTGGGC 1122
721 AAGATATCATCTTATCTATGATTAAGAACTTTGGCTATGCAAAAGCTGGAGCAT 780
|||||
1123 AAGATATCATCTTATCTATGATTAAGAACTTTGGCTATGCAAAAGCTGGAGCAT 1182
781 AGCTGAAGAGCTTTAGACAGCAATCAGACATGATGCTTGGAGAGCAAAAGAAAGA 840
|||||
1183 AGCTGAAGAGCTTTAGACAGCAATGAGACATGATGCTTGGAGAGCAAAAGAAAGA 1242
841 AGCTGAAGAGCTTTAGACAGCAATGAGACATGATGCTTGGAGAGCAAAAGAAAGA 900
1243 AGCTGAAGAGCTTTAGACAGCAATGAGACATGATGCTTGGAGAGCAAAAGAAAGA 1302
901 CACGGCCACATTTCTATTTGCTGGCTGCTTCTATGATCATCATATGCTTGGC 960
1303 TACAGCCATATTTCTATTTGCTGGCTGCTTCTGCTGATCTATGATCATATGCTTGGC 1362
961 TTTCTGATGAGAGCTCTTGGCTGCTTCTGCTGATGATGATGATGATGATGATGAT 1422
1363 CTTCTGATGAGAGCTCTTGGCTGCTTCTGCTGATGATGATGATGATGATGATGAT 1482
1423 TGTATCTTTCTGATTAATTTGAGGCTTTTGTGTGAGCAGGATCTCCAGAGATTGA 1080
1021 TGTCTCTTTCTGATTAATTTGAGGCTTTTGTGTGAGCAGGATCTCCAGAGATTGA 1080
1423 TGTATCTTTCTGATTAATTTGAGGCTTTTGTGTGAGCAGGATCTCCAGAGATTGA 1482
1081 AGCATTTGCAAAAGCAGAGAGAGCTTATGAAATCTTCAAGATTAATTTGCAATTAACC 1140
1483 AGCATTTGCAAAATGCAAGAGAGAGCTTATGAAATCTTCAAGATTAATTTGCAATTAACC 1542
1141 AAGCATTTGCAAGCTATTTGCAAGAGAGCTTAAACAGATTAATTTGCAAGAGAGAG 1200
1543 AAGATTTGCAAGCTATTTGCAAGAGAGCTTAAACAGATTAATTTGCAAGAGAGAG 1602
1201 ATTCAAAATGTTACATTTACCTTCTCGAAAAAGAGATTAGATCTTAAGGCTCT 1260
1603 ATTCAAAATGTTACATTTACCTTCTCGAAAAAGAGATTAGATCTTAAGGCTCT 1662
1261 CAACCTGAAGTTGAGAGTGGGAGACAGTGGCTGTTGGAGACAGTGGCTGGCGGAA 1320

1663 GAACCTGAAGGTTGAGAGAGTGGGAGACAGCTGGCTGGCTGGAGAAACAGTGGCGGAA 1722
1321 GACACAGAGCTGGAGAGCTGATGACAGAGCTTATGACCCCAAGATGATGCTGTAT 1380
1723 GACACAGAGCTGGAGAGCTGATGACAGAGCTTATGACCCCAAGAGAGGAGTGTAGT 1782
1381 TGAATGACAGAGATTTGAGACATTAATGTAAGCATCTTTGGGAAATTTCTGGTGTGT 1440
1783 TGAATGACAGAGATTTGAGACATTAATGTAAGCATCTTTGGGAAATTTCTGGTGTGT 1842
1441 GAGTACAGAGCTGTTGTTTGGCCACAGATAGCTGAAAAATTCGATATGCGCGCA 1500
1843 GAGTACAGAGCTGTTGTTTGGCCACAGATAGCTGAAAAATTCGATATGCGCGCA 1902
1501 AATGTCACCATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1560
1903 AATGTCACCATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1962
1561 CATGAACTACCTATTAATTTGACACTGCTGTTGGAGAGAGAGGAGGAGGAGGAGGAG 1620
1963 CATGAACTACCTATTAATTTGACACTGCTGTTGGAGAGAGAGGAGGAGGAGGAGGAG 2022
1621 TGGACAGAAAGAGAGATTCGCAATGCTGCGGAGGAGGAGGAGGAGGAGGAGGAGGAG 2082
2023 TGGACAGAAAGAGAGATTCGCAATGCTGCGGAGGAGGAGGAGGAGGAGGAGGAGGAG 2140
1681 GCTGATGAGGAG 2142
2083 GCTGATGAGGAG 2202
1741 GATTAAGGAG 1800
2143 GATTAAGGAG 2202
1801 TAATGCGGAGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1860
2203 TAATGCGGAGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 2262
1861 TGAATGCGGAGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1920
2263 TGAATGCGGAGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 2262
1921 TGAATGCGGAGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 2282
2233 TGAATGCGGAGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 2382
1981 GTCCTCAAAAGATTTGAGGCTTCAAGTTTAAATTAAGAAAGATTAATTTGAGGATTAACA 2040
2383 GTCCTCAAAAGATTTGAGGCTTCAAGTTTAAATTAAGAAAGATTAATTTGAGGATTAACA 2442
2041 TGCACCAAGAGGAG 2100
2443 TGCACCAAGAGGAG 2502
2101 TCCAGTTTCTTCTGAGAGATCTGAACTGAACTCAATGAAATGAGGCTTATTTTGTGT 2160
2503 TCCAGTTTCTTCTGAGAGATCTGAACTGAACTCAATGAAATGAGGCTTATTTTGTGT 2562
2161 TGTATATTTTGTGCTATTAATTAAGAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 2220
2563 TGTATATTTTGTGCTATTAATTAAGAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 2622
2221 AAGGATTAAGAGATTTTACCGAGATGAGATCTGAAACAAAGAGAGAGAGAGAGAGAG 2280
2623 AAGGATTAAGAGATTTTACCGAGATGAGATCTGAAACAAAGAGAGAGAGAGAGAGAGAG 2682
2281 CATGTTTCTGATTTGTTTGTAGAGCTTGGAGATTAATTTCTTTATTTTCTCTCA 2340
2683 CATGTTTCTGATTTGTTTGTAGAGCTTGGAGATTAATTTCTTTATTTTCTCTCA 2742
2341 GGGCTTCAATTTGCAAGAGCTGGGAGATCTTCACTAGAGGCTTGTGATGATGATGAT 2400

Db 2743 GGGTTTCACATTGGCAAAAGCTGGAGAGATCCTCACCAGGCGCTCCGATATACATGGTTTT 2802
 QY 2401 CAGATCCATGCTGAGACAGATGTCAGTGGTTTGATGACCCCTAAAGAACCACTGGAGC 2460
 Db 2803 CCGATCCATGCTGACAGAGATGATGATGGTTGATGACCCCTAAAGAACCACTGGAGC 2862
 QY 2461 ATTGACACACGCGCTGGCCAAATGATGGCGCTCAAGTTAAAGGGGCTATAGGTTCCAGGCT 2520
 Db 2863 ATTGACTACACGCGCTCCGCAATGATGCTGCTCAAGTTAAAGGGGCTATAGGTTCCAGGCT 2922
 QY 2521 TGTGTATTAATCCAGAAATATAGCAAACTTTGGGACAGCGCATTTATATATCTTAACTTA 2580
 Db 2923 TGTGTATTAATCCAGAAATATAGCAAACTTTGGGACAGCAATATATATATCTTAACTTA 2982
 QY 2581 TGTGTGCAATTAACACTTTTACTTGTAGCAATTTGACCATATGCAATGCAATAGCAGAGT 2640
 Db 2983 TGTGTGCAACTTAACACTTTTACTTGTAGCAATTTGACCATATGCAATAGCAGAGT 3042
 QY 2641 TGTGAATATGAATATGTTGCTGACAGCACTGAAAAGATGAAGAAAGAGCTGAGAGAGC 2700
 Db 3043 TGTGAATATGAATATGTTGCTGACAGCACTGAAAAGATGAAGAAAGAGCTGAGAGTGC 3102
 QY 2701 TGGGAAGATTGCTACAGAAAGCCATCGAAAACCTTGGAACTGTTGTTCTTGTACTCGGGA 2760
 Db 3103 TGGGAAGATTGCTACAGAAAGCCATCGAAAACCTTGGAACTGTTGTTCTTGTACTCGGGA 3162
 QY 2761 GCAGAAATTGAATATCATGTTATGACACAGAGTTTGCAGAGTACCATCAAACTCTTTGAG 2820
 Db 3163 GCAGAAATTGAATATCATGTTATGACAGAGTTTGCAGAGTACCATCAAACTCTTTGAG 3222
 QY 2821 GAAGACACATCTTCGGGGGCTCATTTTCTATCACCACCAAGCATGATGATATTTTCTTA 2880
 Db 3223 GAAGACACATCTTCGGGGGCTCATTTTCTATCACCACCAAGCATGATGATATTTTCTTA 3282
 QY 2881 TGTGTGCTTTTCGCGTTTGGTTGCTTACTGCTGCAATAGTTCATGAACTTTCAGGA 2940
 Db 3283 TGTGTGCTTTTCGCGTTTGGTTGCTTACTGCTGCAATAGTTCATGAACTTTCAGGA 3342
 QY 2941 TGTGTCTTTGATTTCTGAGCATTTGCTTGGTGGCAATGGCAGTGGGCGAGCTCAATTC 3000
 Db 3343 TGTGTCTTTGATTTTCTGAGCATTTGCTTGGTGGCAATGGCAGTGGGCGAGCTCAATTC 3402
 QY 3001 ATTGTCTCTGACTATGCCAAAGCCAAAGTATCAGACGCCCGCATCATGATCATTTGA 3060
 Db 3403 ATTGTCTCTGACTATGCCAAAGCCAAAGTATCAGACGCCCGCATCATGATCATTTGA 3462
 QY 3061 AAAAAGCCCTGATTTGACAGCTACAGCCCTCAGCGGCTCAAGCCAAATAGTTGGAAG 3120
 Db 3463 AAAAAGCCCTTGAATTTGACAGCTACAGCGGAGGCGCTATATGCCAGACACATTTGGAAG 3522
 QY 3121 AAATGTGACATTTATAGAGTCTGTTCAACTATCCCACTCGACAGACATCCCGTCTG 3180
 Db 3523 AAATGTGACATTTGTTGAAGTTGATTCACACTATCCACCCGAGCCGGAATCCCGTCTG 3582
 QY 3181 CCAGGGGCTGAGCTGAGAGTGAAGAGGCCGACAGCTGGCCCTGTAGTAGAGAGTGG 3240
 Db 3583 TCAGGAGACTAGGCTGAGAGTGAAGAGGCCGACAGCTGGCTGTGTGGAGAGTGG 3642
 QY 3241 CTGTGGAGAGAGCAAGTTGTTCAAGCTCTAGAGCGCTTCTATGAGCCCTTGGCTGGTTC 3300
 Db 3643 CTGTGGAGAGAGCAAGTTGTTCAAGCTCTAGAGCGGTTCTAGAGCCCTTGGCGAGGAA 3702
 QY 3301 AGTGTATTAATTTGATGCAAAAGATTAAGCACTGAAATGTCCAGTGGCTCCGAGCACACT 3360
 Db 3703 AGTGTGTCTTGTATGCAAAAGATTAAGCACTGAAATGTTCAGTGGCTCCGAGCACACT 3762
 QY 3361 GGGCATCTGTTCTCAGAGACCCATCTGTTTACGAGCATTTGGCGAAGCATTTGCCCTA 3420
 Db 3763 GGGCATCTGTTCTCCAGAGACCCATCTGTTTACGAGCATTTGGCGAAGCATTTGCCCTA 3822
 QY 3421 TGGAGACACACCGGGGTGTATCATAGAAAGATTATGACAGCGAGCCAGAGAGCCAA 3480
 Db 3823 TGGAGACACACCGGGGTGTGTACAGAGAGAGATGTTGAGGGCAGAGAAAGAGGCCAA 3882

QY 3481 CATACACACTTCATCCAGACACTCTCTGAAATATCAACACAGAGTAGAGACAAAG 3540
 Db 3883 CATACATCTCTTCACTCAGATCTAGCTGCTTAATTAATATAGCACTAAAGTAGAGACAAAG 3942
 QY 3541 AACCAGCTCTCTGGTGGCCAGAAACAGCCATTTGCCATAGCTGGCTCTTGTAGACA 3600
 Db 3943 AACTCAGCTCTCTGGTGGCCAGAAACAGCCATTTGCCATAGCTGGCTCTTGTAGACA 4002
 QY 3601 GCCCATATTTTGGTTTGGATGAAGCTACATAGCTCTGGATACAGAAAGTAAAGT 3660
 Db 4003 GCCCATATTTTGGTTTGGATGAAGCCAGCTGAGCTCTGGATACAGAAAGTAAAGT 4062
 QY 3661 TGTCCAAAGAGCCCTGGACAAAGCCAGAGAGGCGCCACTGCTGATTTGTATGCCACCG 3720
 Db 4063 TGTCCAAAGAGCCCTGGACAAAGCCAGAGAGGCGCCACTGCTGATTTGTATGCCACCG 4122
 QY 3721 CTGTCCACCATCCAGAAATGACATTTTAATAGTGTGTTTCAAGATGGCAAGTCAAGA 3780
 Db 4123 CTGTCCACCATCCAGAAATGACACTTAATAGTGTGTTTCAAGATGGCAAGTCAAGA 4182
 QY 3781 GCATGGACATCAACACAGCTGCTGGCCAGAAAGCATTTATTTTCCATGGTCAGTGT 3840
 Db 4183 GCATGGACATCAACACAGCTGCTGGCCAGAAAGCATTTATTTTCCATGGTCAGTGT 4242
 QY 3841 CCAGGCTGGAGCAAAAGCGCTAGTGAAGTGGCCATATGAGCTGTTAAATTTTAAAT 3900
 Db 4243 CCAGGCTGGAGCAAAAGCGCCAGTGAAGTGGCCATATGAGCTGTTAAATTTTAAAT 4302
 QY 3901 ATTGTGTTAAACATGCGATTTATTCAAAGTTAAAGTGAAGCATTTACTGAAAACT 3960
 Db 4303 ATT--TGTTAATATGACATTTATTCAAAGTTAA--AGCAAAACATTTACGAAT 4355
 QY 3961 ATGTGAACATCTGTTTAACTTTCTGCAACTGAAAGATCATTCACCAAGTTGAG 4020
 Db 4356 ATGTGAAGATCTGTTTAACTTTCTGCAACTGAAAGATCATTCACCAAGTTGAG 4403
 QY 4021 AOTCTCAGATTTTAAATTTAAAGAACCA--AAGAAACATTTATCGATGAATAAAT 4078
 Db 4404 ----AGAGACTGTAAATTAAGAACAGAGTGAAGATCATCAAGTGAGAGAAAT 4458
 QY 4079 ACTGCTGTTAAATTCATTAATTAATTAAGATTAATCAAGTGAATTTGTTAAAT 4138
 Db 4459 CATAGTTTAAATTCATTAATTAATTAATTAAGATTAATTAAGTGAATTTAAAGTAA 4518
 QY 4139 TGTAAATTTTGTATTTT-----ATTGTAACCTACTCTTGGTGAAGAT 4190
 Db 4519 ATGTGTATTTTGTATTTTATTTTCCATTTGAGCTGTAAGTGAAGTGAAGAT 4578
 QY 4191 TATGAAGTGTAAAGTACTG--AATGTTGAATTAAGTGAAGTGAATTAATTAAGT 4249
 Db 4579 TATGAAGTGTAAAGTGAATTAAGTGAATTTGATTAAGT--TGTAAATTAAGTGA 4635
 QY 4250 ACTTTTAT 4259
 Db 4636 ACTTTCATGT 4645

Search completed: December 9, 2002, 21:10:40
 Job time : 1403 secs

GenCore version 5.1.3
Copyright (c) 1993 - 2002 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: December 9, 2002, 16:58:52 ; Search time 993 Seconds
(without alignments)
9704.230 Million cell updates/sec

Title: US-09-672-725C-1

Perfect score: 4279

Sequence: 1 ggaagcgaggtcggtggtg.....caaaaaaaaaaaaaaa 4279

Scoring table:

IDENTITY NUC
Gapop 10.0 , Gapext 1.0

Searched: 2185239 seqs, 1125999159 residues

Total number of hits satisfying chosen parameters: 4370478

Minimum DB seq length: 0

Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database :

N.Geneseq,101002.*
1: /SIDSL/gcgdata/geneseq/geneseqn-emb1/NA1980.DAT:*
2: /SIDSL/gcgdata/geneseq/geneseqn-emb1/NA1981.DAT:*
3: /SIDSL/gcgdata/geneseq/geneseqn-emb1/NA1982.DAT:*
4: /SIDSL/gcgdata/geneseq/geneseqn-emb1/NA1983.DAT:*
5: /SIDSL/gcgdata/geneseq/geneseqn-emb1/NA1984.DAT:*
6: /SIDSL/gcgdata/geneseq/geneseqn-emb1/NA1985.DAT:*
7: /SIDSL/gcgdata/geneseq/geneseqn-emb1/NA1986.DAT:*
8: /SIDSL/gcgdata/geneseq/geneseqn-emb1/NA1987.DAT:*
9: /SIDSL/gcgdata/geneseq/geneseqn-emb1/NA1988.DAT:*
10: /SIDSL/gcgdata/geneseq/geneseqn-emb1/NA1989.DAT:*
11: /SIDSL/gcgdata/geneseq/geneseqn-emb1/NA1990.DAT:*
12: /SIDSL/gcgdata/geneseq/geneseqn-emb1/NA1991.DAT:*
13: /SIDSL/gcgdata/geneseq/geneseqn-emb1/NA1992.DAT:*
14: /SIDSL/gcgdata/geneseq/geneseqn-emb1/NA1993.DAT:*
15: /SIDSL/gcgdata/geneseq/geneseqn-emb1/NA1994.DAT:*
16: /SIDSL/gcgdata/geneseq/geneseqn-emb1/NA1995.DAT:*
17: /SIDSL/gcgdata/geneseq/geneseqn-emb1/NA1996.DAT:*
18: /SIDSL/gcgdata/geneseq/geneseqn-emb1/NA1997.DAT:*
19: /SIDSL/gcgdata/geneseq/geneseqn-emb1/NA1998.DAT:*
20: /SIDSL/gcgdata/geneseq/geneseqn-emb1/NA1999.DAT:*
21: /SIDSL/gcgdata/geneseq/geneseqn-emb1/NA2000.DAT:*
22: /SIDSL/gcgdata/geneseq/geneseqn-emb1/NA2001A.DAT:*
23: /SIDSL/gcgdata/geneseq/geneseqn-emb1/NA2001B.DAT:*
24: /SIDSL/gcgdata/geneseq/geneseqn-emb1/NA2002.DAT:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	4279	100.0	4279	22	AAD03488
2	4277.4	100.0	4279	22	AAD03504
3	4275.8	99.9	4279	22	AAD03505
4	4272.6	99.9	4279	22	AAD03506
5	4221.8	98.7	4317	22	AAD03489
6	3341.2	78.1	4669	8	AAAT0752
7	3335	77.9	4669	14	AA052726
8	3334.8	77.9	4378	11	AA004522
9	3333.2	77.9	4669	19	AAV32645

10	3333.2	77.9	4669	24	ABK52041	cDNA encoding huma
11	3324.4	77.7	4646	21	AAZ94738	Human ATP binding
12	3324.4	77.7	4646	21	AAD36894	Human mdr1 gene.
13	3324.4	77.7	4646	24	ABL68592	Kidney cancer rela
14	3324.4	77.7	4646	24	ABL68880	Kidney cancer rela
15	3322.6	77.6	4646	22	AAH57442	Human intestine ce
16	3319.6	77.6	4646	22	AAO72872	Human multitrug re
17	3311.8	77.4	4186	22	AAE6127	Cynomolgous monke
18	3311.2	77.4	4195	22	AAE6128	Cynomolgous monke
19	3270.6	76.4	4264	19	AAV66533	Mutated human P-g1
20	3270.6	75.9	4264	19	AAV66534	Human MDR-1 DNA.
21	3247.6	75.7	3988	21	AAZ88973	Hybrid vector pSF-
22	3239.6	75.7	6505	17	AAZ13394	Retroviral vector
23	3226.6	75.4	8630	21	AAZ24042	Human wild-type mu
24	3223.4	75.3	8630	21	AAZ24041	Human BCRP DNA rei
25	3202	74.8	3860	21	AAZ49332	Human G185V mutat
26	3202	74.8	3860	21	ABA94365	Human BCRP DNA rei
27	3200.4	74.8	3860	21	AAZ49333	Human BCRP DNA rei
28	3200.4	74.8	3860	21	ABA94366	Human BCRP DNA rei
29	3190.2	74.6	3840	24	ABL91687	Human BCRP DNA rei
30	2936.2	68.6	4788	21	AAZ49335	Human BCRP DNA rei
31	2936.2	68.6	4788	21	ABA94368	Human BCRP DNA rei
32	2921.4	68.3	4425	21	AAZ52048	Murine polynucleoti
33	28517	68.2	4369	21	AAZ52047	Murine polynucleoti
34	2730.2	63.8	4313	14	AAQ38950	Rat multitrug resi
35	2716.6	63.5	4189	21	AAZ49334	Rat multitrug resi
36	2716.6	63.5	4189	21	ABA94367	Rat multitrug resi
37	2705	63.2	4233	22	AAZ90198	Murine multitrug re
38	2705	63.2	4233	22	AAZ27498	Murine multitrug re
39	2677.4	62.6	4254	22	ABK63517	Mouse BCRP DNA rei
40	2240	52.3	3924	21	AAZ94742	Rat mdr1b2 multidr
41	2240	52.3	3924	21	AAZ88974	Rat mdr1b2 multidr
42	2240	52.3	3924	21	ABN95801	Rat multitrug diff
43	2118	49.5	2726	15	AAQ70907	Human ATP binding
44	2118	49.5	2726	15	AAQ70916	Gene #2299 used to
45	2118	49.5	2726	18	AAZ43322	Multitrug-resistan

ALIGNMENTS

RESULT 1	
ID	AAD03488
AAID03488	standard; cDNA, 4279 BP.
AC	AAID03488:
XX	
DT	13-JUN-2001 (First entry)
XX	
DE	Dog P-glycoprotein (PGP) cDNA #1.
XX	
KW	Dog; P-glycoprotein; PGP; multidrug transporter; MDRL;
KW	drug bioavailability; transgenic animal; genetic model; ss.
OS	Canis familiaris.
XX	
XX	
FT	Key
FT	Location/Qualifiers
FT	17..3862
FT	/tag= a
FT	/Product= "Dog P-glycoprotein (PGP) #1"
PN	WO200123540-A2.
XX	
PD	05-APR-2001.
XX	
PF	28-SEP-2000; 2000WO-US26767.
XX	
FR	28-SEP-1999; 99US-0156510.
XX	
PA	(GENT-) GENTEST CORP.
XX	
PI	Stocker PJ, Steimel-crespi DT, Crespi CL, Reif TC, Patten CJ.
XX	

DR WPI: 2001-235373/24.
XX P-PSDB: AAE00303.
PT New dog P-glycoproteins (Pgp) and their encoding nucleic acids, useful
PT for determining the bioavailability of drugs and for screening for dog
PT Pgp inhibitors -
XX
PS Claim 3: Page 58-63; 111pp; English.
XX The invention relates to dog P-glycoprotein (Pgp) also referred
CC as multidrug transporter (MDR1) and nucleic acids encoding them.
CC The invention also includes fragments and biologically functional
CC variants of dog P-glycoprotein. Pgp and their nucleic acids are
CC useful for determining the bioavailability of drugs and for
CC screening Pgp inhibitors. They are useful for the diagnosis and
CC treatment of conditions characterised by Pgp activity, by
CC reducing or increasing Pgp activity in a cell. Pgp nucleic acids
CC are used as oligonucleotide probes. Complements of Pgp nucleic
CC acids are useful as antisense oligonucleotides, to induce a Pgp
CC 'knockout' phenotype. They are used to prepare a non-human
CC transgenic animal, which are valuable as genetic models for
CC human diseases.
CC The present sequence is dog P-glycoprotein (Pgp) cDNA. This
CC sequence is also referred as Genotype C cDNA. The
CC Pgp enzyme functions as an efflux pump exporting small molecules
CC across the cell membrane. This enzyme is a member of the ABC
CC transporter family.
SQ
SQ Sequence 4279 BP: 1294 A; 834 C; 1008 G; 1143 T; 0 other;
Query Match 100.0%; Score 4279; DB 22; Length 4279;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 4279; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 GGAGCGCCGAGTGGGATGATCTGTAAGAGAGCCGTAAGGGAGTGCAGAGAAACTT 60
DB 1 GGAGCGCCGAGTGGGATGATCTGTAAGAGAGCCGTAAGGGAGTGCAGAGAAACTT 60
QY 61 CTGGAATAATGGGCAAAAAAAGTAAAAAATGAGAAAGAAAGAAAGAAAGAAAGAAAG 120
DB 61 CTGGAATAATGGGCAAAAAAAGTAAAAAATGAGAAAGAAAGAAAGAAAGAAAGAAAG 120
QY 121 CACGTTTGCATTTGCTGATTTCAAAATGGCTTATGATGTTGTTGGGGAG 180
DB 121 CACGTTTGCATTTGCTGATTTCAAAATGGCTTATGATGTTGTTGGGGAG 180
QY 181 AATGGCTGCATTCATCATGAGCTGCACCTCCTCATGATGCTGTTTGGAAACAT 240
DB 181 AATGGCTGCATTCATCATGAGCTGCACCTCCTCATGATGCTGTTTGGAAACAT 240
QY 241 GACAGATAGCTTTGCAAAATGCAAGATTTTCAGAAACAAACCTTCCAGTTAATTAA 300
DB 241 GACAGATAGCTTTGCAAAATGCAAGATTTTCAGAAACAAACCTTCCAGTTAATTAA 300
QY 301 TGAAGATATTACGAACAATACACAATTTTCATCAACCATGAGAGAAATGACAC 360
DB 301 TGAAGATATTACGAACAATACACAATTTTCATCAACCATGAGAGAAATGACAC 360
QY 361 GTATGCTATTATTATACAGTGGATCGTGTGCTGCTGCTGCTGCTTATCATCCAGST 420
DB 361 GTATGCTATTATTATACAGTGGATCGTGTGCTGCTGCTGCTTATCATCCAGST 420
QY 421 TTTCAATTCGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 480
DB 421 TTTCAATTCGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 480
QY 481 TGTATATCATGCGACAGAGATTTGGCTGTTGACGTCAATGAGCTTGGAGTTAATAC 540
DB 481 TGTATATCATGCGACAGAGATTTGGCTGTTGACGTCAATGAGCTTGGAGTTAATAC 540
QY 541 CGGCTCAGACAGATGCTCCAAATCAATGAAGAAATTTGGGACAAATTTGAATGTT 600
DB 541 CGGCTCAGACAGATGCTCCAAATCAATGAAGAAATTTGGGACAAATTTGAATGTT 600

QY 601 CTTTCACTAATAGCAACATTTTTCACCGGTTTATAGTGGGTTTACACGTGTTGAA 660
DB 601 CTTTCACTAATAGCAACATTTTTCACCGGTTTATAGTGGGTTTACACGTGTTGAA 660
QY 661 GCTAACCTTGATTTTGGCCATCAGCCCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 720
DB 661 GCTAACCTTGATTTTGGCCATCAGCCCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 720
QY 721 AAAGATCTATCTTCAATTTATGATTAAGACCTTGGGCTTATGAGAAAGTGAAGCTAT 780
DB 721 AAAGATCTATCTTCAATTTATGATTAAGACCTTGGGCTTATGAGAAAGTGAAGCTAT 780
QY 781 AGCTGAAGAGCTTATGACAGCAATGCAAACTGTATGCTTTGGAGCAAAAGAAAGA 840
DB 781 AGCTGAAGAGCTTATGACAGCAATGCAAACTGTATGCTTTGGAGCAAAAGAAAGA 840
QY 841 ACTTGAAGGTACACAAATTTTGAAGAGTAAAGAAATTTGGATTAAGAAAGCTAT 900
DB 841 ACTTGAAGGTACACAAATTTTGAAGAGTAAAGAAATTTGGATTAAGAAAGCTAT 900
QY 901 CAGGCGCAACATTTCTATGTTGGCTGCTTCTTATGATCTATGATCATATGCTGCTG 960
DB 901 CAGGCGCAACATTTCTATGTTGGCTGCTTCTTATGATCTATGATCATATGCTGCTG 960
QY 961 TTTTGTATGAGACCTTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1020
DB 961 TTTTGTATGAGACCTTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1020
QY 1021 TGTCTCTTTTCTGATTTAATTTGGGCTTTTATGATGAGAGCAATCCCAAGCTATGA 1080
DB 1021 TGTCTCTTTTCTGATTTAATTTGGGCTTTTATGATGAGAGCAATCCCAAGCTATGA 1080
QY 1081 ACATTTTCAAAAGCAAGAGAGAGAGCTTATGAATTTTCAAGATTAATTTACAAATTAAC 1140
DB 1081 ACATTTTCAAAAGCAAGAGAGAGAGCTTATGAATTTTCAAGATTAATTTACAAATTAAC 1140
QY 1141 AAGCATTCAGACTTTTGAAGAGTGAATTAACCAAGATTAATTTAAGGAAATTTGGA 1200
DB 1141 AAGCATTCAGACTTTTGAAGAGTGAATTAACCAAGATTAATTTAAGGAAATTTGGA 1200
QY 1201 ATTCAAAATGTTCACTTCACTTCACTTCACTTCACTTCACTTCACTTCACTTCACT 1260
DB 1201 ATTCAAAATGTTCACTTCACTTCACTTCACTTCACTTCACTTCACTTCACTTCACT 1260
QY 1261 CAACCTGAAGTTTCAAGAGTGGGCAAGAGTGGGCTGTTGGGAAACAGTGGCTCGGAA 1320
DB 1261 CAACCTGAAGTTTCAAGAGTGGGCAAGAGTGGGCTGTTGGGAAACAGTGGCTCGGAA 1320
QY 1321 GAGCAGACCGTGCAGCTGATGAGAGGCTTATGACCCACAGATGCTGCTGAT 1380
DB 1321 GAGCAGACCGTGCAGCTGATGAGAGGCTTATGACCCACAGATGCTGCTGAT 1380
QY 1381 TGAATGACAGACATTAGACATTAATTAAGGCTTTCGAGAAATTTACTGCTGCTG 1440
DB 1381 TGAATGACAGACATTAGACATTAATTAAGGCTTTCGAGAAATTTACTGCTGCTG 1440
QY 1441 GAGTCAGAGACCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1500
DB 1441 GAGTCAGAGACCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1500
QY 1501 AAATGTCACCATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1560
DB 1501 AAATGTCACCATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1560
QY 1561 CATGAACCTTACTATTAATTTGACACTTGTGTTGAGAGAGAGGGGCCACAGTGAATG 1620
DB 1561 CATGAACCTTACTATTAATTTGACACTTGTGTTGAGAGAGAGGGGCCACAGTGAATG 1620
QY 1621 TGGACAGAAACAGAAATGCGCATTTGCTGGGCTTGTGCTGCTGCTGCTGCTGCTGCT 1680
DB 1621 TGGACAGAAACAGAAATGCGCATTTGCTGGGCTTGTGCTGCTGCTGCTGCTGCTGCT 1680

```
QY 1681 GCTGGATGAGGCACTGCTGAGACAGTGAAGTGAAGAGTGGTTCAGGTGGCCCT 1740
DB 1681 GCTGGATGAGGCACTGCTGAGACAGTGAAGTGAAGAGTGGTTCAGGTGGCCCT 1740
QY 1741 GGATAGGCGCAAAAAGCCGAGCTACCATTTGTATAGCTCATGTTTGTATAGTTG 1800
DB 1741 GGATAGGCGCAAAAAGCCGAGCTACCATTTGTATAGCTCATGTTTGTATAGTTG 1800
QY 1801 TAATGCCATGTCATGTTGGTGGTGGATGATGAGTCACTTTGTGAGAAGAAATCATGA 1860
DB 1801 TAATGCCATGTCATGTTGGTGGTGGATGATGAGTCACTTTGTGAGAAGAAATCATGA 1860
QY 1861 TGAAGTCAATGAAGAAGAGGSCATTTACTTCAAACTGTCAATGAGACAGAGAGAA 1920
DB 1861 TGAAGTCAATGAAGAAGAGGSCATTTACTTCAAACTGTCAATGAGACAGAGAGAA 1920
QY 1921 TGAATTTGATTTAGAAAATGCGCATGGTGAATCCAAAAGTGAAGTATGCTTGGAAAT 1980
DB 1921 TGAATTTGATTTAGAAAATGCGCATGGTGAATCCAAAAGTGAAGTATGCTTGGAAAT 1980
QY 1981 GTCTCCAAAAGATTCAGGGTCCAGTTTAAATAAAAGAGATCACTCGAGAGATATCA 2040
DB 1981 GTCTCCAAAAGATTCAGGGTCCAGTTTAAATAAAAGAGATCACTCGAGAGATATCA 2040
QY 2041 TGCACCAAGGCCAAGACAGAAAGCTGGTACAAAAGAGACTTGAATGAGAATGTACC 2100
DB 2041 TGCACCAAGGCCAAGACAGAAAGCTGGTACAAAAGAGACTTGAATGAGAATGTACC 2100
QY 2101 TCCAGTTCTCTTGAGAGATTCGAAAGTGAAGTCAATGAGTGGCTTATTTTGTGGT 2160
DB 2101 TCCAGTTCTCTTGAGAGATTCGAAAGTGAAGTCAATGAGTGGCTTATTTTGTGGT 2160
QY 2161 TGGTATATTTTGTGATTTAAAGAGAGGCTGCAACAGCATTTTCAATATATTTTC 2220
DB 2161 TGGTATATTTTGTGATTTAAAGAGAGGCTGCAACAGCATTTTCAATATATTTTC 2220
QY 2221 AAGGATTAAGGAGATCTTACCCAGATGAGAGATCTGAAACAAAAGACAGAAATAGTAA 2280
DB 2221 AAGGATTAAGGAGATCTTACCCAGATGAGAGATCTGAAACAAAAGACAGAAATAGTAA 2280
QY 2281 CATGTTTCTGATTTGTTCTTACTGTTGAAATATTTCTTTTATACATTTTCCCTCA 2340
DB 2281 CATGTTTCTGATTTGTTCTTACTGTTGAAATATTTCTTTTATACATTTTCCCTCA 2340
QY 2341 GGGCTTCACATTTGGCAAGCTGGGAGATCTCATAAGGGCTTGGATACATGTTTT 2400
DB 2341 GGGCTTCACATTTGGCAAGCTGGGAGATCTCATAAGGGCTTGGATACATGTTTT 2400
QY 2401 CAGATCCATGCTGAGACAGGATGTCAGCTGTTTGAATGACCTTAAAAAACCACTGAGAG 2460
DB 2401 CAGATCCATGCTGAGACAGGATGTCAGCTGTTTGAATGACCTTAAAAAACCACTGAGAG 2460
QY 2461 ATTGACAACAGGCTTCCCATGATGAGGCTCAAGTTAAAGGGGCTATAGTTCCAGGCT 2520
DB 2461 ATTGACAACAGGCTTCCCATGATGAGGCTCAAGTTAAAGGGGCTATAGTTCCAGGCT 2520
QY 2521 TGGTGTATTAACCCAGATTTAGCAAACTTTGGACAGGCATTAATATCTTATATCA 2580
DB 2521 TGGTGTATTAACCCAGATTTAGCAAACTTTGGACAGGCATTAATATCTTATATCA 2580
QY 2581 TGGTGGCAATTAACACTTTTACTTTAGCAATTTGCAATTTGCAATTTGCAAGAGT 2640
DB 2581 TGGTGGCAATTAACACTTTTACTTTAGCAATTTGCAATTTGCAATTTGCAAGAGT 2640
QY 2641 TGGTGAATTAAGAAATGTTTGTGAGACAGACTGAAAGATTAAGAAAGAGCTAGAAAGAGC 2700
DB 2641 TGGTGAATTAAGAAATGTTTGTGAGACAGACTGAAAGATTAAGAAAGAGCTAGAAAGAGC 2700
QY 2701 TGGGAAGATTTGCTACAGAGCCATGAAAACCTTCGAACTGTTTGTGACTCGGGA 2760
DB 2701 TGGGAAGATTTGCTACAGAGCCATGAAAACCTTCGAACTGTTTGTGACTCGGGA 2760
QY 2761 GGAGAGTTTGAATACATGATATGACAGAGTTTGCAGTACATACAGAAACCTTTGAG 2820
DB 2761 GGAGAGTTTGAATACATGATATGACAGAGTTTGCAGTACATACAGAAACCTTTGAG 2820
DB 2761 GGAGAGTTTGAATACATGATATGACAGAGTTTGCAGTACATACAGAAACCTTTGAG 2820
QY 2821 GAAAGACACATCTTGGGGTCTCATTTTATACACCCAGGCATGATATTTTTCCTA 2880
DB 2821 GAAAGACACATCTTGGGGTCTCATTTTATACACCCAGGCATGATATTTTTCCTA 2880
QY 2881 TGGTGGCTGTTCCGGTGGTGGCTTACTGGTGGCAATGAGTTCAATGAACTTTCAGGA 2940
DB 2881 TGGTGGCTGTTCCGGTGGTGGCTTACTGGTGGCAATGAGTTCAATGAACTTTCAGGA 2940
QY 2941 TGTCTTTTGGTATTTTCAAGTATTTGCTTTGGGTCATGAGTGGCCATGAGTGGCCAGTTC 3000
DB 2941 TGTCTTTTGGTATTTTCAAGTATTTGCTTTGGGTCATGAGTGGCCATGAGTGGCCAGTTC 3000
QY 3001 ATTTGCTCTGATCTATGCAAGGCAAAAGATATGAGAGCCAGCCATGATCATATCTTGA 3060
DB 3001 ATTTGCTCTGATCTATGCAAGGCAAAAGATATGAGAGCCAGCCATGATCATATCTTGA 3060
QY 3061 AAAAGCCCTCTGATTTGACAGCTACAGCCCTGACAGGCTCAGCCAAATAGCTTGGAAAG 3120
DB 3061 AAAAGCCCTCTGATTTGACAGCTACAGCCCTGACAGGCTCAGCCAAATAGCTTGGAAAG 3120
QY 3121 AAATGTACATTTAATGAGTCTGTTCAACTATCCACTGACAGACATCCCGTCT 3180
DB 3121 AAATGTACATTTAATGAGTCTGTTCAACTATCCACTGACAGACATCCCGTCT 3180
QY 3181 CCAAGGGCTAGGCTGAGGTTGAAGAAGGCCAGAGCGTGGCCCTGTAAGTACAGTGG 3240
DB 3181 CCAAGGGCTAGGCTGAGGTTGAAGAAGGCCAGAGCGTGGCCCTGTAAGTACAGTGG 3240
QY 3241 CTGTGGGAAGAGACAGTTGTTACGCTCTAGAGGCTTATGAGCCCTTGGTGGTTC 3300
DB 3241 CTGTGGGAAGAGACAGTTGTTACGCTCTAGAGGCTTATGAGCCCTTGGTGGTTC 3300
QY 3301 AGTGTAAATTTAGTGAAGCAAAAGATTAAGCACTGATGAGTGGCTGAGAGACACT 3360
DB 3301 AGTGTAAATTTAGTGAAGCAAAAGATTAAGCACTGATGAGTGGCTGAGAGACACT 3360
QY 3361 GGGCATCTGTGTCAGAGGCCCATCTTGTGACGAGCATTCGCCAGAACTTGGCTA 3420
DB 3361 GGGCATCTGTGTCAGAGGCCCATCTTGTGACGAGCATTCGCCAGAACTTGGCTA 3420
QY 3421 TGGAGACAAGCCGGGTCGTATCATATGAAGATTTATGAGAGCAGCAGAGAGGCCAA 3480
DB 3421 TGGAGACAAGCCGGGTCGTATCATATGAAGATTTATGAGAGCAGCAGAGAGGCCAA 3480
QY 3481 CATACACCACTTCATGAGACACTCCCTGAGAAATACAAACCAAGATGAGAGACAAAG 3540
DB 3481 CATACACCACTTCATGAGACACTCCCTGAGAAATACAAACCAAGATGAGAGACAAAG 3540
QY 3541 AACCCAGCTCTGTGTCAGAGAACAGGCAATGCAATGCTGCGCTTGTATGACA 3600
DB 3541 AACCCAGCTCTGTGTCAGAGAACAGGCAATGCAATGCTGCGCTTGTATGACA 3600
QY 3601 GCTCATATTTTGTGTTGATGATGAGTACATGAGTGGATGAGTACGAAAGTGAAGT 3660
DB 3601 GCTCATATTTTGTGTTGATGATGAGTACATGAGTGGATGAGTACGAAAGTGAAGT 3660
QY 3661 TGTCCAAAGAGCCCTGACAAAGCCAGAGAGCCGACACTGATGATGATGATGATGATG 3720
DB 3661 TGTCCAAAGAGCCCTGACAAAGCCAGAGAGCCGACACTGATGATGATGATGATGATG 3720
QY 3721 CTGTGTCACATCCAGAAATGAGATTTAATGATGATGATGATGATGATGATGATGATG 3780
DB 3721 CTGTGTCACATCCAGAAATGAGATTTAATGATGATGATGATGATGATGATGATGATG 3780
QY 3781 GATGTCACACATGCAAGAGCTGCTGCCCAGAAAGCACTATTTTTCATGATGATGATG 3840
DB 3781 GATGTCACACATGCAAGAGCTGCTGCCCAGAAAGCACTATTTTTCATGATGATGATG 3840
QY 3841 CCAGGCTGAGCAAGGCTGATGAACTGAGTGGCCATATGAGCTTTAAATATTTTAAAT 3900
DB 3841 CCAGGCTGAGCAAGGCTGATGAACTGAGTGGCCATATGAGCTTTAAATATTTTAAAT 3900
```


DB 3841 CCAGGCTGAGCAACGCCCTAGTGAAGCTGGCCATATGACCTGTTAAATTTTTTAAAT 3900
QY 3901 ATTGTGTTAAAGCATGCGATTATATCAAGTTAAAGGTGAGCACTTACTGAAAAACT 3960
DB 3901 ATTGTGTTAAAGCATGCGATTATATCAAGTTAAAGGTGAGCACTTACTGAAAAACT 3960
QY 3961 ATGTGAACTACCTGTTTAAACATTTTCTGCTGCAACTGAGATTCACCAAGTTTCAG 4020
DB 3961 ATGTGAACTACCTGTTTAAACATTTTCTGCTGCAACTGAGATTCACCAAGTTTCAG 4020
QY 4021 AGTCTTCAGATTATATATTAAGCAACCAAGCAACATATCTGATGATTAATATAC 4080
DB 4021 AGTCTTCAGATTATATATTAAGCAACCAAGCAACATATCTGATGATTAATATAC 4080
QY 4081 TGGTGTATATGCAATTTTAAATTAAGATTAATCAAGATTTTGTAAATTAATG 4140
DB 4081 TGGTGTATATGCAATTTTAAATTAAGATTAATCAAGATTTTGTAAATTAATG 4140
QY 4141 TATATATTTTGTATATTTTATTTTGTAACTTACTGCTGAAAGATTTATAGAAGTG 4200
DB 4141 TATATATTTTGTATATTTTATTTTGTAACTTACTGCTGCTGAAAGATTTATAGAAGTG 4200
QY 4201 GTAAAAAGTACTGAATGTTGATTAAGTCAAGCTATATTAATTAACCTTTTATATC 4260
DB 4201 GTAAAAAGTACTGAATGTTGATTAAGTCAAGCTATATTAATTAACCTTTTATATC 4260
QY 4261 AAAAAAAAAAAAAAAAAA 4279
DB 4261 AAAAAAAAAAAAAAAAAA 4279

RESULT 2
AADO3504
ID AADO3504 standard; cDNA; 4279 BP.
XX
AC AADO3504;
XX
DT 13-JUN-2001 (first entry)
XX
DE Dog P-glycoprotein (Pgp) allelic variant (Genotype A) cDNA.
XX
KW Dog: P-glycoprotein allelic variant; multidrug transporter; MDRI;
KM drug bioavailability; transgenic animal; genetic model; ss.
XX
OS Canis familiaris.
XX
FH Key Location/Qualifiers
FT CDS 17..3862
FT /*tag- a
FT /*product- "Dog P-glycoprotein (Pgp) allelic variant
FT (Genotype A) protein"
FT replace (607, C)
FT /*tag- b
XX
XX WO200123540-A2.
XX
XX PD 05-APR-2001.
XX
XX PE 28-SEP-2000; 2000MO-US26767.
XX
XX PR 28-SEP-1999; 9905-0156510.
XX
XX (GENT-) GENEST CORP.
XX
XX Stocker RJ, Steimel-crespi DR, Crespi CL, Reif NC, Patten CJ;
XX
XX WPI: 2001-235373/24.
XX
XX P-PSDB: AAED0308.
XX
XX New dog P-glycoproteins (Pgp) and their encoding nucleic acids, useful
XX for determining the bioavailability of drugs and for screening for dog
XX Pgp inhibitors -

PS Claim 9; Page 85-90; 11pp; English.

XX The invention relates to dog P-glycoprotein (Pgp) also referred
CC as multidrug transporter (MDRI) and nucleic acids encoding them.
CC The invention also includes fragments and biologically functional
CC variants of dog P-glycoprotein. Pgp and their nucleic acids are
CC useful for determining the bioavailability of drugs and for
CC screening Pgp inhibitors. They are useful for the diagnosis and
CC treatment of conditions characterised by Pgp activity/ by
CC reducing or increasing Pgp activity in a cell. Pgp nucleic acids
CC are used as oligonucleotide probes. Complements of Pgp nucleic
CC acids are useful as antisense oligonucleotides, to induce a Pgp
CC "knockout" phenotype. They are used to prepare a non-human
CC transgenic animal, which are valuable as genetic models for
CC human diseases.
CC The present sequence is dog P-glycoprotein (Pgp) allelic variant
CC (Genotype A) cDNA. The Pgp enzyme functions as an efflux pump
CC exporting small molecules across the cell membrane. This enzyme
CC is a member of the ABC transporter family.
XX

SO Sequence 4279 BP; 1295 A; 833 C; 1008 G; 1143 T; 0 other;

Query Match 100.0%; Score 4277.4; DB 22; Length 4279;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 4278; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 GGAGCGCGAGGTGCGATGATCTGAAAGAGCCGTAAGGGAGTGCAGAGAAACTT 60
DB 1 GGAGCGCGAGGTGCGATGATCTGAAAGAGCCGTAAGGGAGTGCAGAGAAACTT 60
QY 61 CTGGAAATATGCGCAAAAAAATGAAAGAAAGAAAGAAAGAAAGAAAGAAAGAAAG 120
DB 61 CTGGAAATATGCGCAAAAAAATGAAAGAAAGAAAGAAAGAAAGAAAGAAAGAAAG 120
QY 121 CACGTTTGCATATGTTTGGCTATTCATTAATGGCTGATAGTGTATGTGTGGGGAC 180
DB 121 CACGTTTGCATATGTTTGGCTATTCATTAATGGCTGATAGTGTATGTGTGGGGAC 180
QY 181 AATGGCTGCGATCATCATGAGTGCACCTCCCTCATGATGCTGATTTTGGAAACAT 240
DB 181 AATGGCTGCGATCATCATGAGTGCACCTCCCTCATGATGCTGATTTTGGAAACAT 240
QY 241 GACGATATGCTTGGCAATATGAGCAAGAAATTCGAAAGAAAGAAAGAAAGAAAGAA 300
DB 241 GACGATATGCTTGGCAATATGAGCAAGAAATTCGAAAGAAAGAAAGAAAGAAAGAA 300
QY 241 GACGATATGCTTGGCAATATGAGCAAGAAATTCGAAAGAAAGAAAGAAAGAAAGAA 300
DB 241 GACGATATGCTTGGCAATATGAGCAAGAAATTCGAAAGAAAGAAAGAAAGAAAGAA 300
QY 301 TGAAGTATATGCAACAAATACACAAATTTTCAACACATCTGAGAGAGAAATGACAC 360
DB 301 TGAAGTATATGCAACAAATACACAAATTTTCAACACATCTGAGAGAGAAATGACAC 360
QY 361 GTATGCTTATTTTAAAGTGGAGTGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 420
DB 361 GTATGCTTATTTTAAAGTGGAGTGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 420
QY 421 TTTCATTTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 480
DB 421 TTTCATTTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 480
QY 481 TGTATATATGCGAGAGATTTGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 540
DB 481 TGTATATATGCGAGAGATTTGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 540
QY 541 CCGGCTCAGAGAGATGCTCCAAATCAATGAAGGAATTTGGGCAAAATTTGGAATGT 600
DB 541 CCGGCTCAGAGAGATGCTCCAAATCAATGAAGGAATTTGGGCAAAATTTGGAATGT 600
QY 601 CTTTCATCTCAATATGCAACATTTTCAACGGTTTATATAGTGGGTTTACAGTGTGAA 660
DB 601 CTTTCATCTCAATATGCAACATTTTCAACGGTTTATATAGTGGGTTTACAGTGTGAA 660
QY 661 GCTAACCTTGTGATTTTGGCCATAGCCCTGTCTTGGACCTTTCAGCCGATCTGGGC 720
DB 661 GCTAACCTTGTGATTTTGGCCATAGCCCTGTCTTGGACCTTTCAGCCGATCTGGGC 720

QY 721 AAGATACATCTTCAATTACGATTAAGAACTCTGGCCATGCAAAAGCTGGACAGT 780
DB 721 AAGATACATCTTCAATTACTGATTAAGAACTCTGGCCATGCAAAAGCTGGACAGT 780
QY 781 ACCTGAAGAAGCTTTAGCAGCAATCAGAACTGATGGCTTTGGAGAGCAAAAGAAAG 840
DB 781 ACCTGAAGAAGCTTTAGCAGCAATCAGAACTGATGGCTTTGGAGAGCAAAAGAAAG 840
QY 841 ACTTGAAGAAGCTTTAGCAAAATTTAGAAAGCTTAAAGGAATTGGATTAAGAAAGCTAT 900
DB 841 ACTTGAAGAAGCTTTAGCAAAATTTAGAAAGCTTAAAGGAATTGGATTAAGAAAGCTAT 900
QY 901 CACGGCCAAACATTTCTATTTGTCGCCCTTTCTATTGATGATGATGATGATGATGATG 960
DB 901 CACGGCCAAACATTTCTATTTGTCGCCCTTTCTATTGATGATGATGATGATGATGATG 960
QY 961 TTTCTGATGAGGACCTTCTGCTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1020
DB 961 TTTCTGATGAGGACCTTCTGCTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1020
QY 1021 TGTCTTTTCTGATTAATTTGAGGCTTTTATGATTTGAGAGGATGCCCAAGCTTAC 1080
DB 1021 TGTCTTTTCTGATTAATTTGAGGCTTTTATGATTTGAGAGGATGCCCAAGCTTAC 1080
QY 1081 AGCATTTGCAAGCAAG 1140
DB 1081 AGCATTTGCAAGCAAG 1140
QY 1141 AAGCATTTGCAAG 1200
DB 1141 AAGCATTTGCAAG 1200
QY 1201 ATTCAAAATTTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1260
DB 1201 ATTCAAAATTTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1260
QY 1261 CAACCTGAAGGTTGAGAGTGGGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1320
DB 1261 CAACCTGAAGGTTGAGAGTGGGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1320
QY 1321 GAGCAG 1380
DB 1321 GAGCAG 1380
QY 1381 TGATGACAG 1440
DB 1381 TGATGACAG 1440
QY 1441 GAGTCAG 1500
DB 1441 GAGTCAG 1500
QY 1501 AAGATGACATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1560
DB 1501 AAGATGACATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1560
QY 1561 CATGAACCTGATTAATTAATTTGATGATGATGATGATGATGATGATGATGATGATG 1620
DB 1561 CATGAACCTGATTAATTAATTTGATGATGATGATGATGATGATGATGATGATGATG 1620
QY 1621 TGGACAGAAACAG 1680
DB 1621 TGGACAGAAACAG 1680
QY 1681 GCTGGATGAGGAG 1740
DB 1681 GCTGGATGAGGAG 1740
QY 1741 GGAATAGGAG 1800
DB 1741 GGAATAGGAG 1800

QY 1801 TAATGCCATGTCATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1860
DB 1801 TAATGCCATGTCATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1860
QY 1861 TGAATCAGTGAAGAAG 1920
DB 1861 TGAATCAGTGAAGAAG 1920
QY 1921 TGAATCAGTGAAGAAG 1980
DB 1921 TGAATCAGTGAAGAAG 1980
QY 1981 GTCACCAAAAGATTCAGAGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 2040
DB 1981 GTCACCAAAAGATTCAGAGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 2040
QY 2041 TGCACCAAAAG 2100
DB 2041 TGCACCAAAAG 2100
QY 2101 TCCAGTTTCTTCTGAG 2160
DB 2101 TCCAGTTTCTTCTGAG 2160
QY 2161 TGGATTAATTTTGTCTATTTAATGAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 2220
DB 2161 TGGATTAATTTTGTCTATTTAATGAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 2220
QY 2221 AAGGATTAATGAAG 2280
DB 2221 AAGGATTAATGAAG 2280
QY 2281 CATGTTTCTGATTTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 2340
DB 2281 CATGTTTCTGATTTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 2340
QY 2341 GGGCTTCAATTTGGCAAAAG 2400
DB 2341 GGGCTTCAATTTGGCAAAAG 2400
QY 2401 CAGATCAGTGTGAG 2460
DB 2401 CAGATCAGTGTGAG 2460
QY 2461 ATTGACAAACAG 2520
DB 2461 ATTGACAAACAG 2520
QY 2521 TGTGTCATTAACCAAGATTAATGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 2580
DB 2521 TGTGTCATTAACCAAGATTAATGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 2580
QY 2581 TGGTTGGCAATTAACAGCTTTTACCTTTAGCAATTTGACCAATGCAATGCAATGCAAT 2640
DB 2581 TGGTTGGCAATTAACAGCTTTTACCTTTAGCAATTTGACCAATGCAATGCAATGCAAT 2640
QY 2641 TGTGGAATGAATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 2700
DB 2641 TGTGGAATGAATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 2700
QY 2701 TGGGAAGATTTGCTACAG 2760
DB 2701 TGGGAAGATTTGCTACAG 2760
QY 2761 GCAGAAATTTGAATACATGATGACAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 2820
DB 2761 GCAGAAATTTGAATACATGATGACAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 2820
QY 2821 GAAAGAGACATCTCGGGGCTGATTTTGTATACAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 2880
DB 2821 GAAAGAGACATCTCGGGGCTGATTTTGTATACAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 2880
QY 2881 TGTGCTGCTGTTCCGCTTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 2940

```

|||||
Db 2881 TGTGGCTGTTCCGGTTTGGTGCTACTGTGGCAAAATGATTCATGAACTTCAGGA 2940
|||
Qy 2941 TGTCTTTTGTATTCACGATATGTCTTGTGGCCATGGGAGGAGGTCAGTTTC 3000
|||||
Db 2941 TGTCTTTTGTATTCACGATATGTCTTGTGGCCATGGGAGGAGGTCAGTTTC 3000
|||
Qy 3001 ATTGTGCTGCTATGATCCAAAGATATCAGACGCCAGCTCATCATGATATTTGA 3060
|||||
Db 3001 ATTGTGCTGCTATGATCCAAAGATATCAGACGCCAGCTCATCATGATATTTGA 3060
|||
Qy 3061 AAAAGCCCTGTGATGACAGCTACAGCCCTCAGCGCCCTCAAGCCCAATATGTTGAAG 3120
|||||
Db 3061 AAAAGCCCTGTGATGACAGCTACAGCCCTCAGCGCCCTCAAGCCCAATATGTTGAAG 3120
|||
Qy 3121 AAATGTACATTTAATGAGGTCGTGTTCAACTATCCACTGACACAGACATCCCGTCT 3180
|||||
Db 3121 AAATGTACATTTAATGAGGTCGTGTTCAACTATCCACTGACACAGACATCCCGTCT 3180
|||
Qy 3181 CCAGGGGCTGAGCTCGAGGTGAAGAGGCCAGAGCGCGCCCTGTAGTACAGTGG 3240
|||||
Db 3181 CCAGGGGCTGAGCTCGAGGTGAAGAGGCCAGAGCGCGCCCTGTAGTACAGTGG 3240
|||
Qy 3241 CTGTGGGAAGGACAGTTGTGATGCTCTAGAGCGCTTCTATGACCCCTTGCTGCTTC 3300
|||||
Db 3241 CTGTGGGAAGGACAGTTGTGATGCTCTAGAGCGCTTCTATGACCCCTTGCTGCTTC 3300
|||
Qy 3301 AGTGTAAATGATGGCAAGATAAAGACCTGAATGTCAGTGGCTCCGACACCT 3360
|||||
Db 3301 AGTGTAAATGATGGCAAGATAAAGACCTGAATGTCAGTGGCTCCGACACCT 3360
|||
Qy 3361 GGGCATGCTGCTCGAGGCCATCTGTTGATGCTGACGATTTGCCGAACATTTGCCCTA 3420
|||||
Db 3361 GGGCATGCTGCTCGAGGCCATCTGTTGATGCTGACGATTTGCCGAACATTTGCCCTA 3420
|||
Qy 3421 TGGAGACAACAGCCGGTCTGATCATCATGAAGATTTAGGAGCAGCCAGAGAGCCAA 3480
|||||
Db 3421 TGGAGACAACAGCCGGTCTGATCATCATGAAGATTTAGGAGCAGCCAGAGAGCCAA 3480
|||
Qy 3481 CATACACCTTCATCGAGACACTCCCTGAGAAATACAAACAGAGTAGAGACAAAG 3540
|||||
Db 3481 CATACACCTTCATCGAGACACTCCCTGAGAAATACAAACAGAGTAGAGACAAAG 3540
|||
Qy 3541 AACCCAGCTCTCTGTTGGCCAGAAACAGCGCATTTCCATACCTGCGCTCTTTAGACA 3600
|||||
Db 3541 AACCCAGCTCTCTGTTGGCCAGAAACAGCGCATTTCCATACCTGCGCTCTTTAGACA 3600
|||
Qy 3601 GCCTCATATTTGCTTTGGATGAAGCTACATCAGCTGATCAGAAAGTGAAGGT 3660
|||||
Db 3601 GCCTCATATTTGCTTTGGATGAAGCTACATCAGCTGATCAGAAAGTGAAGGT 3660
|||
Qy 3661 TGTCCAGAAGACCCCTGGACAAAGCCAGAGAGCCGACCTCATTTGATGATGCCACCG 3720
|||||
Db 3661 TGTCCAGAAGACCCCTGGACAAAGCCAGAGAGCCGACCTCATTTGATGATGCCACCG 3720
|||
Qy 3721 CTGTCCACCATCCAGATCGAATTTAATAGTGTTCAGAAATGCGAAAGTCAAGGA 3780
|||||
Db 3721 CTGTCCACCATCCAGATCGAATTTAATAGTGTTCAGAAATGCGAAAGTCAAGGA 3780
|||
Qy 3781 GCATGCGACATCATCAGAGCTGTGCGCCAGAAAGGCATCTATTTTCATGCTGATGT 3840
|||||
Db 3781 GCATGCGACATCATCAGAGCTGTGCGCCAGAAAGGCATCTATTTTCATGCTGATGT 3840
|||
Qy 3841 CCAGGCTGAGACAAAGGCTATGATGATGCGCATATGATGCTTAAATATTTTAAAT 3900
|||||
Db 3841 CCAGGCTGAGACAAAGGCTATGATGATGCGCATATGATGCTTAAATATTTTAAAT 3900
|||
Qy 3901 ATTGTGTTAAAGATGATTTAATCAAGTTAAAGGTGAGACACTTCTGAAAAAAT 3960
|||||
Db 3901 ATTGTGTTAAAGATGATTTAATCAAGTTAAAGGTGAGACACTTCTGAAAAAAT 3960
|||
Qy 3961 ATGTGAACCTACCTGTTTAAATTTCTTGTCGAAGTGAAGATCATTCACCAAGTTCAG 4020
|||||

```

```

Db 3961 ATGTGAACCTACCTGTTTAACTTTCTTGCTGCAACTGAGATCATTTCCACCAAGTTCAG 4020
Qy 4021 AGCTTTCAGATTTTAAATTAAGAACCAAGAACCAATTAATCTGATGATTAATTAATAC 4080
|||
Db 4021 AGCTTTCAGATTTTAAATTAAGAACCAAGAACCAATTAATCTGATGATTAATTAATAC 4080
|||
Qy 4081 TGTGTTAAATTCATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 4140
|||||
Db 4081 TGTGTTAAATTCATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 4140
|||
Qy 4141 TATTAATTTTGTATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 4200
|||||
Db 4141 TATTAATTTTGTATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 4200
|||
Qy 4201 GTAAAAAGTACGATGTTTGAATTAAGTACGCTATTAATTAATTAATTAATTAATTC 4260
|||||
Db 4201 GTAAAAAGTACGATGTTTGAATTAAGTACGCTATTAATTAATTAATTAATTAATTC 4260
|||
Qy 4261 AAAAAAAAAAAAAAAAAAAAAA 4279
|||||
Db 4261 AAAAAAAAAAAAAAAAAAAAAA 4279

```

RESULT 3

AAD03505 standard; cDNA; 4279 BP.

AAD03505;

13-JUN-2001 (first entry)

Dog P-glycoprotein (PGP) allelic variant (Genotype B) cDNA.

Dog; P-glycoprotein allelic variant; PGP; multidrug transporter; MDRL; drug bioavailability; transgenic animal; genetic model; ss.

Canis familiaris.

Location/Qualifiers

Key /tag= a Location/Qualifiers

FT /product= "Dog P-glycoprotein (PGP) allelic variant"

FT (Genotype B) protein"

FT replace (91, T)

FT /*tag= b

FT replace (607, C)

FT /*tag= c

PN W0200123540-A2.

PD 05-APR-2001.

XX 28-SEP-2000; 2000WO-US267.

XX 28-SEP-1999; 9905-0156510.

XX (GENT-) GENTEST CORP.

XX Stocker PJ, Steinel-crespi DR, Crespi CL, Reif TC, Patten CJ;

XX WPI; 2001-235373/24.

XX P-PSDB; AAE00309.

XX New dog P-glycoproteins (PGP) and their encoding nucleic acids, useful

XX for determining the bioavailability of drugs and for screening for dog

XX PGP inhibitors -

XX Claim 9; Page 93-99; 11pp; English.

XX The invention relates to dog P-glycoprotein (PGP) also referred

XX as multidrug transporter (MDRL) and nucleic acids encoding them.

XX The invention also includes fragments and biologically functional

XX variants of dog P-glycoprotein. PGP and their nucleic acids are

Db 1861 TGAACATGAAAGAGAGGCGATTACTTCAAACTGTCAACATGACAGACAGAGAAA 1920
OY 1921 TGAATTTGATTAAGAAATGCCACTGGTGAATCCAAAGGAAAGTATGCTTGGAAAT 1980
OY 1921 TGAATTTGATTAAGAAATGCCACTGGTGAATCCAAAGGAAAGTATGCTTGGAAAT 1980
OY 1981 GTCTCCAAAAGATGAGGCTCAGTTTAAATAAAAAGAGATCAACCTGCAGAGTATACA 2040
OY 1981 GTCTCCAAAAGATGAGGCTCAGTTTAAATAAAAAGAGATCAACCTGCAGAGTATACA 2040
Db 1981 GTCTCCAAAAGATGAGGCTCAGTTTAAATAAAAAGAGATCAACCTGCAGAGTATACA 2040
OY 2041 TGCACCAAGAGGCGAAGACAGAACTTGTACAAAAGAGACTTGAATGAGATGACC 2100
OY 2041 TGCACCAAGAGGCGAAGACAGAACTTGTACAAAAGAGACTTGAATGAGATGACC 2100
Db 2041 TGCACCAAGAGGCGAAGACAGAACTTGTACAAAAGAGACTTGAATGAGATGACC 2100
OY 2101 TCCAGTTTCTGTGAGAGATCTGAGAGCTGAACTCAACTGAATGCGCTTATTTTGTGGT 2160
OY 2101 TCCAGTTTCTGTGAGAGATCTGAGAGCTGAACTCAACTGAATGCGCTTATTTTGTGGT 2160
Db 2101 TCCAGTTTCTGTGAGAGATCTGAGAGCTGAACTCAACTGAATGCGCTTATTTTGTGGT 2160
OY 2161 TGGTATATTTTGTCTATTTAAACGAGGCTGCAACAGCATTTTCAATATATTTTC 2220
OY 2161 TGGTATATTTTGTCTATTTAAACGAGGCTGCAACAGCATTTTCAATATATTTTC 2220
Db 2161 TGGTATATTTTGTCTATTTAAACGAGGCTGCAACAGCATTTTCAATATATTTTC 2220
OY 2221 AAGGATATAGGAGATCTTACCCAGATAGAGATCTGAAACAAACAGACAGATATGTA 2280
OY 2221 AAGGATATAGGAGATCTTACCCAGATAGAGATCTGAAACAAACAGACAGATATGTA 2280
Db 2221 AAGGATATAGGAGATCTTACCCAGATAGAGATCTGAAACAAACAGACAGATATGTA 2280
OY 2281 CATGTTTCTGTATTTGTCTTACTCTTGGAAATTTATTTTATTTTACTTTTCTCTCA 2340
OY 2281 CATGTTTCTGTATTTGTCTTACTCTTGGAAATTTATTTTATTTTACTTTTCTCTCA 2340
Db 2281 CATGTTTCTGTATTTGTCTTACTCTTGGAAATTTATTTTATTTTACTTTTCTCTCA 2340
OY 2341 GGGCTTCATTTGGCAAGCGTGGGAGATCTCACTAAGCGGCTTGCATACATGTTTT 2400
OY 2341 GGGCTTCATTTGGCAAGCGTGGGAGATCTCACTAAGCGGCTTGCATACATGTTTT 2400
Db 2341 GGGCTTCATTTGGCAAGCGTGGGAGATCTCACTAAGCGGCTTGCATACATGTTTT 2400
OY 2401 CAGATCCATGCTGACAGAGATGTCAGCTGGTTTGAATGACCCCTTAAACACACATGAGAC 2460
OY 2401 CAGATCCATGCTGACAGAGATGTCAGCTGGTTTGAATGACCCCTTAAACACACATGAGAC 2460
Db 2401 CAGATCCATGCTGACAGAGATGTCAGCTGGTTTGAATGACCCCTTAAACACACATGAGAC 2460
OY 2461 ATTTGCAACCCAGGCTTGGCAATGATGGGCTCAAGTTAAAGGGCTATAGTTCCAGGCT 2520
OY 2461 ATTTGCAACCCAGGCTTGGCAATGATGGGCTCAAGTTAAAGGGCTATAGTTCCAGGCT 2520
Db 2461 ATTTGCAACCCAGGCTTGGCAATGATGGGCTCAAGTTAAAGGGCTATAGTTCCAGGCT 2520
OY 2521 TGGTGTATTTACCCAGAAATATAGCAATCTTGGAGACAGGATTTATTTTATCTTAATCTA 2580
OY 2521 TGGTGTATTTACCCAGAAATATAGCAATCTTGGAGACAGGATTTATTTTATCTTAATCTA 2580
Db 2521 TGGTGTATTTACCCAGAAATATAGCAATCTTGGAGACAGGATTTATTTTATCTTAATCTA 2580
OY 2581 TGGTGTGCAATTTACACTTTTACTCTTAGCAATTTGACCAATCTTGAATGAGAGAT 2640
OY 2581 TGGTGTGCAATTTACACTTTTACTCTTAGCAATTTGACCAATCTTGAATGAGAGAT 2640
Db 2581 TGGTGTGCAATTTACACTTTTACTCTTAGCAATTTGACCAATCTTGAATGAGAGAT 2640
OY 2641 TGTGAAATGAAATGTTGTCTGACAGACACTGAAATGAATGAAGAGCTAGAGAGAGC 2700
OY 2641 TGTGAAATGAAATGTTGTCTGACAGACACTGAAATGAATGAAGAGCTAGAGAGAGC 2700
Db 2641 TGTGAAATGAAATGTTGTCTGACAGACACTGAAATGAATGAAGAGCTAGAGAGAGC 2700
OY 2701 TGGGAATATGTTGACAGAAAGCTTGAAGAACTTGCAGAACTGTTGTTCTTGTACTCGGA 2760
OY 2701 TGGGAATATGTTGACAGAAAGCTTGAAGAACTTGCAGAACTGTTGTTCTTGTACTCGGA 2760
Db 2701 TGGGAATATGTTGACAGAAAGCTTGAAGAACTTGCAGAACTGTTGTTCTTGTACTCGGA 2760
OY 2761 GCAGAAATTTGAATACATGTATGACAGAGTTTGAAGAGACATTCAGAAATCTTTGAG 2820
OY 2761 GCAGAAATTTGAATACATGTATGACAGAGTTTGAAGAGACATTCAGAAATCTTTGAG 2820
Db 2761 GCAGAAATTTGAATACATGTATGACAGAGTTTGAAGAGACATTCAGAAATCTTTGAG 2820
OY 2821 GAAAGCACACATCTTGGGGGCTCATTTTCTATCACCCAGGCAATGATGATTTTCTTA 2880
OY 2821 GAAAGCACACATCTTGGGGGCTCATTTTCTATCACCCAGGCAATGATGATTTTCTTA 2880
Db 2821 GAAAGCACACATCTTGGGGGCTCATTTTCTATCACCCAGGCAATGATGATTTTCTTA 2880
OY 2881 TGTGCTGTTTCCGGTTTGGTGTCTTACTTGTGCGCAATGAGTTTCTTA 2940
OY 2881 TGTGCTGTTTCCGGTTTGGTGTCTTACTTGTGCGCAATGAGTTTCTTA 2940
Db 2881 TGTGCTGTTTCCGGTTTGGTGTCTTACTTGTGCGCAATGAGTTTCTTA 2940
OY 2941 TGTCTTTTGTGATCTAGCTATTTGCTTGTGAGCATGAGAGTGGGCGAGTCACTTC 3000
OY 2941 TGTCTTTTGTGATCTAGCTATTTGCTTGTGAGCATGAGAGTGGGCGAGTCACTTC 3000
Db 2941 TGTCTTTTGTGATCTAGCTATTTGCTTGTGAGCATGAGAGTGGGCGAGTCACTTC 3000

OY 3001 ATTTGCTCTGATATGCCAAAGCCAAAGTATGACAGCCACAGCTCATATGATATGTA 3060
OY 3001 ATTTGCTCTGATATGCCAAAGCCAAAGTATGACAGCCACAGCTCATATGATATGTA 3060
Db 3001 ATTTGCTCTGATATGCCAAAGCCAAAGTATGACAGCCACAGCTCATATGATATGTA 3060
OY 3061 AAAAGCCCTGTATTTGACAGCTACAGCCCTCAGGCTTCAAGCCAAATACGTTGGAAG 3120
OY 3061 AAAAGCCCTGTATTTGACAGCTACAGCCCTCAGGCTTCAAGCCAAATACGTTGGAAG 3120
Db 3061 AAAAGCCCTGTATTTGACAGCTACAGCCCTCAGGCTTCAAGCCAAATACGTTGGAAG 3120
OY 3121 AAATGTACATTTAATAGAGTGTGTTCATATCCACTGACCTGACCAAGCATCCCCGCT 3180
OY 3121 AAATGTACATTTAATAGAGTGTGTTCATATCCACTGACCTGACCAAGCATCCCCGCT 3180
Db 3121 AAATGTACATTTAATAGAGTGTGTTCATATCCACTGACCTGACCAAGCATCCCCGCT 3180
OY 3181 CCAGGCGCTGAGCTCGAGTGAAGAAAGGCGAGAGCGTGGCCCTCTAGTAGCAGTGG 3240
OY 3181 CCAGGCGCTGAGCTCGAGTGAAGAAAGGCGAGAGCGTGGCCCTCTAGTAGCAGTGG 3240
Db 3181 CCAGGCGCTGAGCTCGAGTGAAGAAAGGCGAGAGCGTGGCCCTCTAGTAGCAGTGG 3240
OY 3241 CTGAGGAGAGACAGATTTTCAAGCTCTAGAGCGCTTCTATGACCCCTTGGGCTGTTTC 3300
OY 3241 CTGAGGAGAGACAGATTTTCAAGCTCTAGAGCGCTTCTATGACCCCTTGGGCTGTTTC 3300
Db 3241 CTGAGGAGAGACAGATTTTCAAGCTCTAGAGCGCTTCTATGACCCCTTGGGCTGTTTC 3300
OY 3301 AGTGTATTTATGCGCAAAAGATTAAGACCTGAAATGTCAGATGCTCGAGACACACT 3360
OY 3301 AGTGTATTTATGCGCAAAAGATTAAGACCTGAAATGTCAGATGCTCGAGACACACT 3360
Db 3301 AGTGTATTTATGCGCAAAAGATTAAGACCTGAAATGTCAGATGCTCGAGACACACT 3360
OY 3361 GGGCATGCTGTCTAGAGAGCCCATCTGTTTGAATGACATGATGCGGAGAACATTTGCTTA 3420
OY 3361 GGGCATGCTGTCTAGAGAGCCCATCTGTTTGAATGACATGATGCGGAGAACATTTGCTTA 3420
Db 3361 GGGCATGCTGTCTAGAGAGCCCATCTGTTTGAATGACATGATGCGGAGAACATTTGCTTA 3420
OY 3421 TGGAGCAACAGCCGGGTGCTGATCAATGAAAGATATGACAGGCGCAAGAGGCGCA 3480
OY 3421 TGGAGCAACAGCCGGGTGCTGATCAATGAAAGATATGACAGGCGCAAGAGGCGCA 3480
Db 3421 TGGAGCAACAGCCGGGTGCTGATCAATGAAAGATATGACAGGCGCAAGAGGCGCA 3480
OY 3481 CATACACACTTCATGAGAGACCTCCTGAGAAATATACACACAGATAGAGACAAAG 3540
OY 3481 CATACACACTTCATGAGAGACCTCCTGAGAAATATACACACAGATAGAGACAAAG 3540
Db 3481 CATACACACTTCATGAGAGACCTCCTGAGAAATATACACACAGATAGAGACAAAG 3540
OY 3541 AACCCAGCTCTGTGGTGCAGAAACAGCGCATTTGCCATGCTGCGCTTGTATGACA 3600
OY 3541 AACCCAGCTCTGTGGTGCAGAAACAGCGCATTTGCCATGCTGCGCTTGTATGACA 3600
Db 3541 AACCCAGCTCTGTGGTGCAGAAACAGCGCATTTGCCATGCTGCGCTTGTATGACA 3600
OY 3601 GCGCATATTTTCTTTTGGATGAAGCTATCATCATGCTTGTGATACAAAGTGAAGAGT 3660
OY 3601 GCGCATATTTTCTTTTGGATGAAGCTATCATCATGCTTGTGATACAAAGTGAAGAGT 3660
Db 3601 GCGCATATTTTCTTTTGGATGAAGCTATCATCATGCTTGTGATACAAAGTGAAGAGT 3660
OY 3661 TGTCCAAAGAGCCCTGAGACAAAGCCAGAGAAAGCGGCACTGATGATGCGCCACCG 3720
OY 3661 TGTCCAAAGAGCCCTGAGACAAAGCCAGAGAAAGCGGCACTGATGATGCGCCACCG 3720
Db 3661 TGTCCAAAGAGCCCTGAGACAAAGCCAGAGAAAGCGGCACTGATGATGCGCCACCG 3720
OY 3721 CTGTGCAACATCCAGAAATGCAATTTAATAGTGTGTTTCGAAATGCGCAAGTCAAGA 3780
OY 3721 CTGTGCAACATCCAGAAATGCAATTTAATAGTGTGTTTCGAAATGCGCAAGTCAAGA 3780
Db 3721 CTGTGCAACATCCAGAAATGCAATTTAATAGTGTGTTTCGAAATGCGCAAGTCAAGA 3780
OY 3781 GCATGCGACACATTAACAGTGTGCGCCAGAAAGGCTATTTTCCATGCTGCTAGT 3840
OY 3781 GCATGCGACACATTAACAGTGTGCGCCAGAAAGGCTATTTTCCATGCTGCTAGT 3840
Db 3781 GCATGCGACACATTAACAGTGTGCGCCAGAAAGGCTATTTTCCATGCTGCTAGT 3840
OY 3841 CCAGGCTGAGCAAGGCGTATGTAAGTGTGCGCATATGAGCTGTAAATATTTTAAAT 3900
OY 3841 CCAGGCTGAGCAAGGCGTATGTAAGTGTGCGCATATGAGCTGTAAATATTTTAAAT 3900
Db 3841 CCAGGCTGAGCAAGGCGTATGTAAGTGTGCGCATATGAGCTGTAAATATTTTAAAT 3900
OY 3901 AATTTGTAAATGATGCGCATTTTAAATCAAGTTTAAAGGTGAGCACTTACTGAAAACT 3960
OY 3901 AATTTGTAAATGATGCGCATTTTAAATCAAGTTTAAAGGTGAGCACTTACTGAAAACT 3960
Db 3901 AATTTGTAAATGATGCGCATTTTAAATCAAGTTTAAAGGTGAGCACTTACTGAAAACT 3960
OY 3961 ATGTGAACTACCTGTTTAAATTTTCTGTCACATGAAATGATATTCACCAAGTTGAG 4020
OY 3961 ATGTGAACTACCTGTTTAAATTTTCTGTCACATGAAATGATATTCACCAAGTTGAG 4020
Db 3961 ATGTGAACTACCTGTTTAAATTTTCTGTCACATGAAATGATATTCACCAAGTTGAG 4020
OY 4021 AGTCTTCAGATTTTAAATTAAGAACCAAAACATTAATCTGATGGAATTAATATAC 4080
OY 4021 AGTCTTCAGATTTTAAATTAAGAACCAAAACATTAATCTGATGGAATTAATATAC 4080
Db 4021 AGTCTTCAGATTTTAAATTAAGAACCAAAACATTAATCTGATGGAATTAATATAC 4080

RESULT 4
AAD03506
ID AAD03506 standard; cDNA; 4279 BP.
XX

Canis familiaris.
OS
XX
KM
MDRL; drug bioavailability; transgenic animal; genetic model; ss.
dog; glycoprotein allelic variant; PGP; multidrug transporter;

WO200123540-A2

26-SEP-2000; 2000MO-US26767.

28-SEP-1999; 9905-0156510.

(GENT-) GENTEST CORP.

..... cresspl LU, Kelt JC, Patten CJ;
X

P-PSDB; AAEE00310.

for determining the bioavailability of drugs and for screening for dog PGP inhibitors -

claim 5, Page 102-101; 111pp; English.

as multidrug transporter (MDR1) and nucleic acids encoding them. The invention also includes fragments and biologically functional variants of dog p-glycoprotein. Pgp and their nucleic acids are useful for determining the bioavailability of drugs and for screening PGP inhibitors. They are useful for the diagnosis and

frequency 42/5 BF; 1290 A; 833 C; 1009 G; 1141 T; 0 other;

Query Match	99.9%;	Score 4272.6;	DB 22;	Length 4279;
Best Local Similarity	99.9%;	Pred. No. 0;		
Matches 4275: Conservative	0.0;	Mismatches 4		

γ , units $0; \infty$; gaps $0;$

1 GGAGCGCGAGGTCGGGATGATCCTGAAGGAGGAGCCGTAAGGGAGTCCAGAGAAACTTT 60
1 GGAGCGCGAGGTCGGGATGATCCTGAAGGAGGAGCCGTAAGGGAGTCCAGAGAAACTTT 60

120
 121
 122
 123
 124
 125
 126
 127
 128
 129
 130
 131
 132
 133
 134
 135
 136
 137
 138
 139
 140
 141
 142
 143
 144
 145
 146
 147
 148
 149
 150
 151
 152
 153
 154
 155
 156
 157
 158
 159
 160
 161
 162
 163
 164
 165
 166
 167
 168
 169
 170
 171
 172
 173
 174
 175
 176
 177
 178
 179
 180
 181
 182
 183
 184
 185
 186
 187
 188
 189
 190
 191
 192
 193
 194
 195
 196
 197
 198
 199
 200
 201
 202
 203
 204
 205
 206
 207
 208
 209
 210
 211
 212
 213
 214
 215
 216
 217
 218
 219
 220
 221
 222
 223
 224
 225
 226
 227
 228
 229
 230
 231
 232
 233
 234
 235
 236
 237
 238
 239
 240
 241
 242
 243
 244
 245
 246
 247
 248
 249
 250
 251
 252
 253
 254
 255
 256
 257
 258
 259
 260
 261
 262
 263
 264
 265
 266
 267
 268
 269
 270
 271
 272
 273
 274
 275
 276
 277
 278
 279
 280
 281
 282
 283
 284
 285
 286
 287
 288
 289
 290
 291
 292
 293
 294
 295
 296
 297
 298
 299
 300
 301
 302
 303
 304
 305
 306
 307
 308
 309
 310
 311
 312
 313
 314
 315
 316
 317
 318
 319
 320
 321
 322
 323
 324
 325
 326
 327
 328
 329
 330
 331
 332
 333
 334
 335
 336
 337
 338
 339
 340
 341
 342
 343
 344
 345
 346
 347
 348
 349
 350
 351
 352
 353
 354
 355
 356
 357
 358
 359
 360
 361
 362
 363
 364
 365
 366
 367
 368
 369
 370
 371
 372
 373
 374
 375
 376
 377
 378
 379
 380
 381
 382
 383
 384
 385
 386
 387
 388
 389
 390
 391
 392
 393
 394
 395
 396
 397
 398
 399
 400
 401
 402
 403
 404
 405
 406
 407
 408
 409
 410
 411
 412
 413
 414
 415
 416
 417
 418
 419
 420
 421
 422
 423
 424
 425
 426
 427
 428
 429
 430
 431
 432
 433
 434
 435
 436
 437
 438
 439
 440
 441
 442
 443
 444
 445
 446
 447
 448
 449
 450
 451
 452
 453
 454
 455
 456
 457
 458
 459
 460
 461
 462
 463
 464
 465
 466
 467
 468
 469
 470
 471
 472
 473
 474
 475
 476
 477
 478
 479
 480
 481
 482
 483
 484
 485
 486
 487
 488
 489
 490
 491
 492
 493
 494
 495
 496
 497
 498
 499
 500
 501
 502
 503
 504
 505
 506
 507
 508
 509
 510
 511
 512
 513
 514
 515
 516
 517
 518
 519
 520
 521
 522
 523
 524
 525
 526
 527
 528
 529
 530
 531
 532
 533
 534
 535
 536
 537
 538
 539
 540
 541
 542
 543
 544
 545
 546
 547
 548
 549
 550
 551
 552
 553
 554
 555
 556
 557
 558
 559
 560
 561
 562
 563
 564
 565
 566
 567
 568
 569
 570
 571
 572
 573
 574
 575
 576
 577
 578
 579
 580
 581
 582
 583
 584
 585
 586
 587
 588
 589
 590
 591
 592
 593
 594
 595
 596
 597
 598
 599
 600
 601
 602
 603
 604
 605
 606
 607
 608
 609
 610
 611
 612
 613
 614
 615
 616
 617
 618
 619
 620
 621
 622
 623
 624
 625
 626
 627
 628
 629
 630
 631

...GAGGAAGAAGAAACCACTGT CAG 120

180

221 CACGCTTCCGCTATTCAAATGGCTGATAGGTTGTATATCTTGGTGGGAC 180

201 TTTTGGAAACAT 240

5'-TGGGACATCTGTTTTCGATGAGCCTGCACCTCCCTCATGATGCTGTTTTCGAACAT-3' 240

211 CACGCGHAGCCTTGCAAAATGCAGGAATTTCAAGAAACAACAACCTTTCCAGTTATAATTAA 300

211 CACGCHAGCCTTGGAAATTCAGAGAACAACCTTTCACGTTATAATTAA 300

|||||ATGAAACAATACCAACATTTTCATCAACCATCTGGAGGAGGAATGACCAC 360

501 TGGAGCAATACACACACATTTCATCAACCATCTGGAGGAGGAATGACCAC 360

501 GATACCAATACAGTGGGATCGGTGCTGGCGTGTGGTGGCTGCTTACATCCAAGT 420

561 GAAAGCCTAATTAACAGTGGGATCGGCTGCTGGCGTGTGCTGGCTGCTTACATCCAGT 420

721
11081166166CAGGACAGATCTCAAATTAGAAACAATTTTCA 480

421 TCTATCTGGTGGCAGCAGATCTCAAAATTAGAAACAATTCTTCA 480

481 TGCATCAATGGACAGGACATTGGCTGTTTGACGTGCATGACGTTGGGAGCTTAACAC 540

481 TGCATCAATGCGACGAGATTGGCTGGTTTGACCTGCATGACGTTGGGAGCTTAACAC 540

341 CCGGTCACAGACGATGTCCTCAAATCAATGAAGGAATTGGCGACAAATTCGAATGTT 600

341 CCGGCTCACAGACGATGTCCTCCAAATCAATGAGGAATTGGCGACAAAATTGGAATGTT 600

601 CTTTCACTCAATAGCAACATTTTTCACCGGTTTATAGTGGGGTTTACACGCTGTTGGA 660

601 CTTTCAATCATAGCACACATTTTTCACCGGTTTATAGTGGGTTTACACGTGGTTGGA 660

561 GCYACCCCTTGATTTTGGCCATCAGCCCTGTTCTTGACTTTCAGCCGCCATCTGGG 720

661 GCACCCCTTGATTTGGCCATCAGCCCTGTTCTTGGACTTTCAGCCGCATCTGGC 720

121 AAAGATACCTATCTTCATTTACTGATAAAGAACTCTTGGCCTATGCAAAAGCTGGAGCAGT 780

121 AAAGATACCTATCTTCATTCTACTGATAAGAACTCTTGGCTATGCAAAGCTGCACCGCT 790

81 AGCTGAGAGTCTTAGCAGCAATCAGAACTGTGATTGCCCTTTGAGGACCAACACCA 940

781 AGCTGAGAGTCTTAGCAGCAATCAGACTGTGATTCCTTTGAGGACAAAGCAACA 840

QY	841	ACCTGAAAGGACAAACAAAATTAGTAAGAAAGCTAAAGAAATGGGATTAAGAAAGCTAT	900
Db	841	ACTTGAAAAGTACAAACAAAATTAGTAAGAAAGCTAAAGAAATGGGATTAAGAAAGCTAT	900
QY	901	CACGGCCACATTTCTATATGGTGGCGGTTCTTATATGATCATATGATCATATATGCTTGCG	960
Db	901	CACGGCCACATTTCTATATGGTGGCGGTTCTTATATGATCATATGATCATATATGCTTGCG	960
QY	961	TTTCTGATATGGAGACTCTTGCTGCTCCGAGTAATTTCTATATGGACAACTACTATCAC	1020
Db	961	TTTCTGATATGGAGACTCTTGCTGCTCCGAGTAATTTCTATATGGACAACTACTATCAC	1020
QY	1021	TGCTCTTTTCTCTGATTAATTTGGGGGCTTTAGATTTGGACAGGACATCCCAAGCATTTGA	1080
Db	1021	TGCTCTTTTCTCTGATTAATTTGGGGGCTTTAGATTTGGACAGGACATCCCAAGCATTTGA	1080
QY	1081	AGCATTTGCAACGCAAGAGAGCAGCTTATGAATCTTCAAGTAATTTGACATAATAAC	1140
Db	1081	AGCATTTGCAACGCAAGAGAGCAGCTTATGAATCTTCAAGTAATTTGACATAATAAC	1140
QY	1141	AAGCATTTGACGCTTTTCGACAGAGGACATTAACACAGATTAATTAAGGAAATTTGGA	1200
Db	1141	AAGCATTTGACGCTTTTCGACAGAGGACATTAACACAGATTAATTAAGGAAATTTGGA	1200
QY	1201	ATTCAAAAATGTTCACTTCAGTTACCCCTTCGAAAAGAGTTAAGTCTTAAGGGCTCT	1260
Db	1201	ATTCAAAAATGTTCACTTCAGTTACCCCTTCGAAAAGAGTTAAGTCTTAAGGGGCTCT	1260
QY	1261	CAACCTGAAGGTTCAAGATGGGAGACAGATGGCGCTGTTGGGAACAAGTGGCTCGGGAA	1320
Db	1261	CAACCTGAAGGTTCAAGATGGGAGACAGATGGCGCTGTTGGGAACAAGTGGCTCGGGAA	1320
QY	1321	GAGCAGCAGCCGTGACGTGATGGAGAGGCTATGACCCCAAGATGGCATGCTGTAT	1380
Db	1321	GAGCAGCAGCCGTGACGTGATGGAGAGGCTATGACCCCAAGATGGCATGCTGTAT	1380
QY	1381	TGATGGACAGGACATTAGACACATAAATGTAAAGCATCTTCGGGAAATTAAGTGGTGGT	1440
Db	1381	TGATGGACAGGACATTAGACACATAAATGTAAAGCATCTTCGGGAAATTAAGTGGTGGT	1440
QY	1441	GAGTCAGAGGCGTGTGTTTGGCCACACAGATAGCTGAAAACATTGCGTATGGCCGCA	1500
Db	1441	GAGTCAGAGGCGTGTGTTTGGCCACACAGATAGCTGAAAACATTGCGTATGGCCGCA	1500
QY	1501	AAATGTCACATGATGATGATTTGAAAAGCTGTTAAGGAAGCCAAATGCCATATATTTAT	1560
Db	1501	AAATGTCACATGATGATGATTTGAAAAGCTGTTAAGGAAGCCAAATGCCATATATTTAT	1560
QY	1561	CATGAACTACTATTAATTTGACACTCTGTTGGAGAGAGAGGGGCCACAGCTGAGTGG	1620
Db	1561	CATGAACTACTATTAATTTGACACTCTGTTGGAGAGAGAGGGGCCACAGCTGAGTGG	1620
QY	1621	TGGACAGAAACAGAGATTCGCCATTTGCTGGGGCCCTGGTTGCAACCCCAAGATTTCTCT	1680
Db	1621	TGGACAGAAACAGAGATTCGCCATTTGCTGGGGCCCTGGTTGCAACCCCAAGATTTCTCT	1680
QY	1681	GCGGATGAGGCAAGCTCAGCTCTGGACACTGAAAAGTGAAGCAAGTGGTTCAAGTGGCCCT	1740
Db	1681	GCGGATGAGGCAAGCTCAGCTCTGGACACTGAAAAGTGAAGCAAGTGGTTCAAGTGGCCCT	1740
QY	1741	GGATTAAGCCAGAAAAGGCCGAGACATTAATTTGATACCTCATGCTTGTCTACAGTTGC	1800
Db	1741	GGATTAAGCCAGAAAAGGCCGAGACATTAATTTGATACCTCATGCTTGTCTACAGTTGC	1800
QY	1801	TAAATGCGATGATTTGCTGTTTTGATGATGAGTCAATTTGGAGAAAGGAATTCATGA	1860
Db	1801	TAAATGCGATGATTTGCTGTTTTGATGATGAGTCAATTTGGAGAAAGGAATTCATGA	1860
QY	1861	TGAATCTATGAAGAAGAGGCAATTAACCTCAAACTGTCAAAATGCAAGCAAGAGAAA	1920
Db	1861	TGAATCTATGAAGAAGAGGCAATTAACCTCAAACTGTCAAAATGCAAGCAAGAGAAA	1920

QY	1921	TGAATTCAGTTTACAAAATGCACATGGTGAATCCAAAAGTGAAGTGAATGATCCTTGGAAAT	1980
Db	1921	TGAATTCAGTTTACAAAATGCACATGGTGAATCCAAAAGTGAAGTGAATGATCCTTGGAAAT	1980
QY	1981	GTCCTCAAAAGATTCAGGGTCAGTTTAATAAAAGAGATCAACTCGCAGAGATATACA	2040
Db	1981	GTCCTCAAAAGATTCAGGGTCAGTTTAATAAAAGAGATCAACTCGCAGAGATATACA	2040
QY	2041	TGCACCAAGGCCAAGACAAAGCTTGGTACAAAAGAGACTTGATGAGATGTACC	2100
Db	2041	TGCACCAAGGCCAAGACAAAGCTTGGTACAAAAGAGACTTGATGAGATGTACC	2100
QY	2101	TCCAGTTTCCTCTGAGAGATTCGAAAGCTGAACCTGAACGTATTTTGTGT	2160
Db	2101	TCCAGTTTCCTCTGAGAGATTCGAAAGCTGAACCTGAACGTATTTTGTGT	2160
QY	2161	TGGTATTTTGTGATTTATTAACGAGAGGCTCGAACAGCATTTTCAATATATTTTC	2220
Db	2161	TGGTATTTTGTGATTTATTAACGAGAGGCTCGAACAGCATTTTCAATATATTTTC	2220
QY	2221	AAGGATTTATAGGATCTTTACCCGAGTGAATCCGAACAAAGACAGATAATGTAA	2280
Db	2221	AAGGATTTATAGGATCTTTACCCGAGTGAATCCGAACAAAGACAGATAATGTAA	2280
QY	2281	CATGTTTTCGTATTTGTTCTTCACTGCTTGGATTAATTTCTTTATTAACATTTTCCCA	2340
Db	2281	CATGTTTTCGTATTTGTTCTTCACTGCTTGGATTAATTTCTTTATTAACATTTTCCCA	2340
QY	2341	GGGCTTCACATTTTGGCAAAAGCTGGGAGATTCCTCAAGGGCTCGATTCATAGTTTTT	2400
Db	2341	GGGCTTCACATTTTGGCAAAAGCTGGGAGATTCCTCAAGGGCTCGATTCATAGTTTTT	2400
QY	2401	CAGATTCATGCTGAGACAGAGATGTACGTGGTTTGTATGACCTTAAAAACACCACTCGAGC	2460
Db	2401	CAGATTCATGCTGAGACAGAGATGTACGTGGTTTGTATGACCTTAAAAACACCACTCGAGC	2460
QY	2461	ATTGACAAACAGGCTTCCCAATGATGGGGTCAAGTTAAAGGGCTATAGTTCCAGGCT	2520
Db	2461	ATTGACAAACAGGCTTCCCAATGATGGGGTCAAGTTAAAGGGCTATAGTTCCAGGCT	2520
QY	2521	TGCTGTCAATCCAGATATATAGCAATCTTTGGACAGCATTTATATCTCTTAATCTA	2580
Db	2521	TGCTGTCAATCCAGATATATAGCAATCTTTGGACAGCATTTATATCTCTTAATCTA	2580
QY	2581	TGGTTGGCAATTAACATTTTACCTTACGTATGCAATTTGTAACCATATGTGCAATAGCAGAGT	2640
Db	2581	TGGTTGGCAATTAACATTTTACCTTACGTATGCAATTTGTAACCATATGTGCAATAGCAGAGT	2640
QY	2641	TGTTGAAATGAATAATGTTGTCTGGACAAAGCACTGAAAGATTAAGAAAGAGCTTAAGAGAC	2700
Db	2641	TGTTGAAATGAATAATGTTGTCTGGACAAAGCACTGAAAGATTAAGAAAGAGCTTAAGAGAC	2700
QY	2701	TGGGAAGATTGCTACAGAAGCATCGAAAATTCGAACTGTGTTCTTGAATCGGGA	2760
Db	2701	TGGGAAGATTGCTACAGAAGCATCGAAAATTCGAACTGTGTTCTTGAATCGGGA	2760
QY	2761	GCAGAGTTTGAATACATGTATGACAGAGTTTGCAGATACATACAGAAACTTTTGAAG	2820
Db	2761	GCAGAGTTTGAATACATGTATGACAGAGTTTGCAGATACATACAGAAACTTTTGAAG	2820
QY	2821	GAAAGCACATCTTGGGGCTCATTTTCTATATCCAGGCATGTATATTTTCTCTA	2880
Db	2821	GAAAGCACATCTTGGGGCTCATTTTCTATATCCAGGCATGTATATTTTCTCTA	2880
QY	2881	TGCTGGCTGTTCCGGTTTGTGTCCTACTTGTGTGCAATGATTCATGAACCTTTCAGA	2940
Db	2881	TGCTGGCTGTTCCGGTTTGTGTCCTACTTGTGTGCAATGATTCATGAACCTTTCAGA	2940
QY	2941	TGTTCTTTTGGTATTCACCTATTTCTTTTGGTCATAGGAGGGGCAAGTTCAGTTC	3000
Db	2941	TGTTCTTTTGGTATTCACCTATTTCTTTTGGTCATAGGAGGGGCAAGTTCAGTTC	3000
QY	3001	ATTTCGTCGACATATGCCAAAGCCAAAGATATACGAGCCACGTCATCATGATCATTTGA	3060

```

Db 3001 ATTTGCTCCGTGACTATGCAAAAGCATATGACGAGCCACGTCATCATGTCAATTA 3060
QY 3061 AAAAGCCCTGATTTGACCTACAGCCCTCACGSCCTCAAGCCAAATACGTTGGAAG 3120
Db 3061 AAAAGCCCTGATTTGACCTACAGCCCTCACGSCCTCAAGCCAAATACGTTGGAAG 3120
QY 3121 AATATGACATTTAATGAGTCTGTTCACATATCCCACTCGCCAGACATATCCCGTGT 3180
Db 3121 AATATGACATTTAATGAGTCTGTTCACATATCCCACTCGCCAGACATATCCCGTGT 3180
QY 3181 CCAGGGGCTGAGCCTCGAGGTGAAGAAAGGCGCACGCTGCGCTCTGTAGTAGCAGTG 3240
Db 3181 CCAGGGGCTGAGCCTCGAGGTGAAGAAAGGCGCACGCTGCGCTCTGTAGTAGCAGTG 3240
QY 3241 CTGTGGGAAGAGCAGCAGTGTTCAGCTCCTAGAGCGCTCTGTATGACCCCTTGCTG 3300
Db 3241 CTGTGGGAAGAGCAGCAGTGTTCAGCTCCTAGAGCGCTCTGTATGACCCCTTGCTG 3300
QY 3301 AGTCTAATTTGATGGCAAGAGATTAAGCAGCTGAATGCTCAGTGGCTCCGAGCACCT 3360
Db 3301 AGTCTAATTTGATGGCAAGAGATTAAGCAGCTGAATGCTCAGTGGCTCCGAGCACCT 3360
QY 3361 GGGCATCTGTCTCAGAGAGCCCATCCTGTTGATGTCAGCACTTGGCCGAGAACTTGCCTA 3420
Db 3361 GGGCATCTGTCTCAGAGAGCCCATCCTGTTGATGTCAGCACTTGGCCGAGAACTTGCCTA 3420
QY 3421 TGGAGCAACAGCGCGGTGTATCATATGAAGATTAATGACAGCAGCCAGAGAGGCCAA 3480
Db 3421 TGGAGCAACAGCGCGGTGTATCATATGAAGATTAATGACAGCAGCCAGAGAGGCCAA 3480
QY 3481 CATACACCACTTCATGAGACACTCCCTGAGAAATACAAACACAGAGTAGAGACAAAG 3540
Db 3481 CATACACCACTTCATGAGACACTCCCTGAGAAATACAAACACAGAGTAGAGACAAAG 3540
QY 3541 AACCAGCTCTGTGTGCCCAAGAAACAGCCGATTCGAGCTCGGCTCTGTTAGACA 3600
Db 3541 AACCAGCTCTGTGTGCCCAAGAAACAGCCGATTCGAGCTCGGCTCTGTTAGACA 3600
QY 3601 GCGTCATATTTTGGTTTGGATGAAAGTACATCAGCTCTGAGATACAGAAAGTGAAGAGT 3660
Db 3601 GCGTCATATTTTGGTTTGGATGAAAGTACATCAGCTCTGAGATACAGAAAGTGAAGAGT 3660
QY 3661 TGTCCAAAGACCTTGACAAAGCCAGAGAGGCGCACCTGATTTGATGCCACCG 3720
Db 3661 TGTCCAAAGACCTTGACAAAGCCAGAGAGGCGCACCTGATTTGATGCCACCG 3720
QY 3721 CTGTCCACCATCCAGATGCAAGATTAATGATGATGTTTCAGAATGGCAAGTCAAGGA 3780
Db 3721 CTGTCCACCATCCAGATGCAAGATTAATGATGATGTTTCAGAATGGCAAGTCAAGGA 3780
QY 3781 GCATGGCAGCATCAACAGCTGCTGCGCCAGAAAGGCAATGATGTTTCAAGTCAAGGA 3840
Db 3781 GCATGGCAGCATCAACAGCTGCTGCGCCAGAAAGGCAATGATGTTTCAAGTCAAGGA 3840
QY 3841 CCAAGGCTGAGCAAGCGGTAGTGAAGTGGCCATATGAGCTGTTAAATTTTTTAAAT 3900
Db 3841 CCAAGGCTGAGCAAGCGGTAGTGAAGTGGCCATATGAGCTGTTAAATTTTTTAAAT 3900
QY 3901 ATTTTGTGTTAAACATGAGATTAAATCAAGTTAAAGTAGAGTACTGAGAAACT 3960
Db 3901 ATTTTGTGTTAAACATGAGATTAAATCAAGTTAAAGTAGAGTACTGAGAAACT 3960
QY 3961 ATGTAACATACCTGTTTAAACATTTCTTGGTGCAGATGAGATCATCCCAAGTTGAG 4020
Db 3961 ATGTAACATACCTGTTTAAACATTTCTTGGTGCAGATGAGATCATCCCAAGTTGAG 4020
QY 4021 AGCTTTCAGATTTTAATTAAGGAACCAAAAGAAACATTAATGGAATTAATAC 4080
Db 4021 AGCTTTCAGATTTTAATTAAGGAACCAAAAGAAACATTAATGGAATTAATAC 4080
QY 4081 TGTGTTAATGCAATTAATAATTAATAGAGTAATCAAAAGTAGATTTTGAATAATG 4140

```

```

Db 4081 TGTGTTAATGCAATTAATAATTAATAGAGTAATCAAAAGTAGATTTTGAATAATG 4140
QY 4141 TATTAATTTTGGTTTAAATTTTAAATTTGTAACCTACTGCTTGTCTGAAGATTAAGAGTG 4200
Db 4141 TATTAATTTTGGTTTAAATTTTAAATTTGTAACCTACTGCTTGTCTGAAGATTAAGAGTG 4200
QY 4201 GTAAAAAGTACGAAATGTTTGAATTAAGTGTACCTATTAATTAACCTTTATATC 4260
Db 4201 GTAAAAAGTACGAAATGTTTGAATTAAGTGTACCTATTAATTAACCTTTATATC 4260
QY 4261 AAAAAAAAAAAAAAAAAA 4279
Db 4261 AAAAAAAAAAAAAAAAAA 4279

RESULT 5
AAD03489
ID AAD03489 standard; cDNA; 4317 BP.
XX
AC AAD03489;
XX
DT 13-JUN-2001 (first entry)
XX
DE Dog P-glycoprotein (PGP) cDNA #2.
XX
KW Dog; P-glycoprotein; PGP; multidrug transporter; MDRL;
KW drug bioavailability; transgenic animal; genetic model; ss.
OS Canis familiaris.
XX
FH Key
FT Location/Qualifiers
FT 70..3912
FT CDS
FT /tag= a
FT /product= "Dog P-glycoprotein (PGP) #2"

WO200123540-A2.
05-APR-2001.
28-SEP-2000; 2000WO-US26767.
28-SEP-1999; 99US-0156510.
(GENT-) GENTEST CORP.
Stocker PJ, Steimel-crespi DT, Crespi CL, Reif TC, Patten CJ;
WPI: 2001-235373/24.
P-PSDB; MAE00304.

New dog P-glycoproteins (PGP) and their encoding nucleic acids, useful
for determining the bioavailability of drugs and for screening for dog
PGP inhibitors -
PS Claim 1; Page 66-72; 11pp; English.
XX
CC The invention relates to dog P-glycoprotein (PGP) also referred
CC as multidrug transporter (MDRL) and nucleic acids encoding them.
CC The invention also includes fragments and biologically functional
CC variants of dog P-glycoprotein. PGP and their nucleic acids are
CC useful for determining the bioavailability of drugs and for
CC screening PGP inhibitors. They are useful for the diagnosis and
CC treatment of conditions characterised by PGP activity, by
CC reducing or increasing PGP activity in a cell. PGP nucleic acids
CC are used as oligonucleotide probes. Complements of PGP nucleic
CC 'knockout' phenotype. They are used to prepare a non-human
CC transgenic animal, which are valuable as genetic models for
CC human diseases.
CC The present sequence is dog P-glycoprotein (PGP) cDNA. The
CC PGP enzyme functions as an efflux pump exporting small molecules
CC across the cell membrane. This enzyme is a member of the ABC
CC transporter family.

```



```
QY 2103 CAGTTTCCTTGGAGAGATTCGAAAGCGAACCTGAACTGAAAGGCTTATTTTGGGTG 2162
|||||
Db 2153 CAGTTTCCTTGGAGAGATTCGAAAGCGTGAACCTGAACTGAAAGGCTTATTTTGGGTG 2212
|||||
QY 2163 GTATATTTTGTCTATTTATTAAGAGGCGCGACACAGACATTTTCATAATATTTTCAA 2222
|||||
Db 2213 GTATATTTTGTCTATTTATTAAGAGGCGCGACACAGACATTTTCATAATATTTTCAA 2272
|||||
QY 2223 GGATTTATAGGATCTTTTACCCGAGATGAGATTCCTGAACAAAGACAGAAATAGTAA 2282
|||||
Db 2273 GGATTTATAGGATCTTTTACCCGAGATGAGATTCCTGAACAAAGACAGAAATAGTAA 2332
|||||
QY 2283 TGTTTCTGTATTTGTTCAGTCCCTGGAAATTTATTTCTTTTATATACATTTTCCACAG 2342
|||||
Db 2333 TGTTTCTGTATTTGTTCAGTCCCTGGAAATTTATTTCTTTTATATACATTTTCCACAG 2392
|||||
QY 2343 GCTTCACATTTTGGCAAGCTGGGGAGATCCCACTAAAGCGGCTTGATACATAGTTTCA 2402
|||||
Db 2393 GCTTCACATTTTGGCAAGCTGGGGAGATCCCACTAAAGCGGCTTGATACATAGTTTCA 2452
|||||
QY 2403 GATCCATGCTGAGACAGATGTCAGCTGCTTGTATGACCCCTAAACACCACTGGAGCAT 2462
|||||
Db 2453 GATCCATGCTGAGACAGATGTCAGCTGCTTGTATGACCCCTAAACACCACTGGAGCAT 2512
|||||
QY 2463 TGACAAACAGGCTTGGCAATGATGCGGCTCAAGTTAAAGGGGCTATAGGTTCCAGGCTG 2522
|||||
Db 2513 TGACAAACAGGCTTGGCAATGATGCGGCTCAAGTTAAAGGGGCTATAGGTTCCAGGCTG 2572
|||||
QY 2523 CTGTCAATTTCCGAAATATAGCAAACTTTGGGACAGGCAATTTATATCCCTAATCATAG 2582
|||||
Db 2573 CTGTCAATTTCCGAAATATAGCAAACTTTGGGACAGGCAATTTATATCCCTAATCATAG 2632
|||||
QY 2583 GTTGGCAATTTACACTTTTACTTTAGCAATTTGATACCATCATTTGATAGCAGAGAGTTG 2642
|||||
Db 2633 GTTGGCAATTTACACTTTTACTTTAGCAATTTGATACCATCATTTGATAGCAGAGAGTTG 2692
|||||
QY 2643 TTGAATGAATAATGTTGTCTGACAGAGACTGAAAGATTAAGAAAGCTAGAAAGAGCTG 2702
|||||
Db 2693 TTGAATGAATAATGTTGTCTGACAGAGACTGAAAGATTAAGAAAGCTAGAAAGAGCTG 2752
|||||
QY 2703 GGAAGATTTGCTACAGAAAGCATCGAAAACTTCGAACTGTGTTCTTTGACTGGGAGC 2762
|||||
Db 2753 GGAAGATTTGCTACAGAAAGCATCGAAAACTTCGAACTGTGTTCTTTGACTGGGAGC 2812
|||||
QY 2763 AGAAGTTGAATACATGTAAGCAGAGAGTTGCAAGTACCATACAGAACTCTTTGAGA 2822
|||||
Db 2813 AGAAGTTGAATACATGTAAGCAGAGAGTTGCAAGTACCATACAGAACTCTTTGAGA 2872
|||||
QY 2823 AAGCACACATCTCGGGGCTCATTTTCTATACCCAGGCAATGATATTTTCTATG 2882
|||||
Db 2873 AAGCACACATCTCGGGGCTCATTTTCTATACCCAGGCAATGATATTTTCTATG 2932
|||||
QY 2883 CTGGCTGTTTTCGGGTTTGGTGTCTACCTTGGTGGCAAAATAGTTCATGAACTTCAGATG 2942
|||||
Db 2933 CTGGCTGTTTTCGGGTTTGGTGTCTACCTTGGTGGCAAAATAGTTCATGAACTTCAGATG 2992
|||||
QY 2943 TTTCTTTTGGTATTTCTCAGCTATTTGTTTGGTGCATGGCAGTGGGGCAGTCACTCAT 3002
|||||
Db 2993 TTTCTTTTGGTATTTCTCAGCTATTTGTTTGGTGCATGGCAGTGGGGCAGTCACTCAT 3052
|||||
QY 3003 TTGCTCTGACTATGGCAAAAGCAAGTATCAGCAGGCCACGTCATCATGATCATGATAA 3062
|||||
Db 3053 TTGCTCTGACTATGGCAAAAGCAAGTATCAGCAGGCCACGTCATCATGATCATGATAA 3112
|||||
QY 3063 AAAACCCCTGATTTGACACTACAGCCCTCAGCGGCTCAAGCAAAATGCTTGAAGGAA 3122
|||||
Db 3113 AAAACCCCTGATTTGACACTACAGCCCTCAGCGGCTCAAGCAAAATGCTTGAAGGAA 3172
|||||
QY 3123 ATGTGACATTTATAGAGGCTGCTTCAACTATCCCACTGACCAAGACATCCCGCTGCTC 3182
|||||
Db 3173 ATGTGACATTTATAGAGGCTGCTTCAACTATCCCACTGACCAAGACATCCCGCTGCTC 3232
|||||

QY 3183 AGGGGCTGAGCCCTCGAGGTGAAGAGAGGCCAGACGCTGAGCCCTGTAAGTAGAGGCT 3242
|||||
Db 3233 AGGGGCTGAGCCCTCGAGGTGAAGAGAGGCCAGACGCTGAGCCCTGTAAGTAGAGGCT 3292
|||||
QY 3243 GTGGGAAGACACATGTTGTACGCTCTAGAGGCTTTCATATACCCCTTGGCTGGTTCA 3302
|||||
Db 3293 GTGGGAAGACACATGTTGTACGCTCTAGAGGCTTTCATATACCCCTTGGCTGGTTCA 3352
|||||
QY 3303 TGCTAATTTATGCAAGAGATTAAGACCTGATGTCCAGTGGCTCCGACACACTGG 3362
|||||
Db 3353 TGCTAATTTATGCAAGAGATTAAGACCTGATGTCCAGTGGCTCCGACACACTGG 3412
|||||
QY 3363 GCATCGTGTCTGAGAGGCCCATCCCTGTTTACAGCAGCATTTGCCAGAAATCTGCTATG 3422
|||||
Db 3413 GCATCGTGTCTGAGAGGCCCATCCCTGTTTACAGCAGCATTTGCCAGAAATCTGCTATG 3472
|||||
QY 3423 GAGACAAAGCCGGGCTGATACATGAAAGATTTTGAAGGACAGCCAAAGAGGCCAAGA 3482
|||||
Db 3473 GAGACAAAGCCGGGCTGATACATGAAAGATTTTGAAGGACAGCCAAAGAGGCCAAGA 3532
|||||
QY 3483 TACACACTTCATGAGACACTCCCTGAGAAATACAAACCGAGTATGAGACAAAGGAA 3542
|||||
Db 3533 TACACACTTCATGAGACACTCCCTGAGAAATACAAACCGAGTATGAGACAAAGGAA 3592
|||||
QY 3543 CCCAGCTCTGTGTTGGCCAGAAACAGCCCATTTGCCATTAAGCTTGCCTCTTTTGAACAGC 3602
|||||
Db 3593 CCCAGCTCTGTGTTGGCCAGAAACAGCCCATTTGCCATTAAGCTTGCCTCTTTTGAACAGC 3652
|||||
QY 3603 CTGATATTTTGTGTTTGGATGATACATCACTGCTGATGATGATGATGATGATGATGATG 3662
|||||
Db 3653 CTGATATTTTGTGTTTGGATGATACATCACTGCTGATGATGATGATGATGATGATGATG 3712
|||||
QY 3663 TCCAAAGACCCCTGGACAAAGCCAGAGAGAGCCGACCTGCTATTTGATGCTCCACGCT 3722
|||||
Db 3713 TCCAAAGACCCCTGGACAAAGCCAGAGAGAGCCGACCTGCTATTTGATGCTCCACGCT 3772
|||||
QY 3723 TGTCCACATCCAGAAATCAGATTTAATAGTGTGTTTTCAGATGAGCAAACTCAAGAGC 3782
|||||
Db 3773 TGTCCACATCCAGAAATCAGATTTAATAGTGTGTTTTCAGATGAGCAAACTCAAGAGC 3832
|||||
QY 3783 ATGAGCACATCAAAAGCTGTGGCCAGAAAGGCACTCAATTTTTCATGCTGATGCTC 3842
|||||
Db 3833 ATGAGCACATCAAAAGCTGTGGCCAGAAAGGCACTCAATTTTTCATGATGATGCTC 3892
|||||
QY 3843 AGGCTGAGCAAGGCTGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 3902
|||||
Db 3893 AGGCTGAGCAAGGCTGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 3952
|||||
QY 3903 TTTGTTTAAACATGAGCATTTTATCAAAAGTTAAAGGTGAGCACTTACTGGAAAACTAT 3962
|||||
Db 3953 TTTGTTTAAACATGAGCATTTTATCAAAAGTTAAAGGTGAGCACTTACTGGAAAACTAT 4012
|||||
QY 3963 GTAGAACCTACCTGTTTAACTTTCTGCTGCACTGAAATATTCACCAAGTTCCAGAG 4022
|||||
Db 4013 GTAGAACCTACCTGTTTAACTTTCTGCTGCACTGAAATATTCACCAAGTTCCAGAG 4072
|||||
QY 4023 TCTTCAGATTTTATATTAAGAACCAAAAGAACATTTATGATGATGATGATGATGAT 4082
|||||
Db 4073 TCTTCAGATTTTATATTAAGAACCAAAAGAACATTTATGATGATGATGATGATGAT 4132
|||||
QY 4083 GTGTATATTGCATTTATTAATTAAGATTAAGATTAAGATTAAGATTAAGATTAAGATTA 4142
|||||
Db 4133 GTGTATATTGCATTTATTAATTAAGATTAAGATTAAGATTAAGATTAAGATTAAGATTA 4192
|||||
QY 4143 TATTTTGTATATTTATTTATTTATTTATTTATTTATTTATTTATTTATTTATTTATTT 4202
|||||
Db 4193 TATTTTGTATATTTATTTATTTATTTATTTATTTATTTATTTATTTATTTATTTATTT 4252
|||||
QY 4203 AAAAAGTACTGAATGTTGAATTAAGAGTACTGATTAATTAATTAATTAATTAATTA 4262
|||||
Db 4253 AAAAAGTACTGAATGTTGAATTAAGAGTACTGATTAATTAATTAATTAATTAATTA 4312
|||||
QY 4263 AAAAA 4267
```

Db 4313 AAAA 4317

RESULT 6

AA070752 standard; cDNA; 4669 BP.

AA070752;

21-MAY-1991 (first entry)

Sequence of human multi-drug resistance-1 (mdr1) cDNA from clones lambda-HDR10.5 and 104.

Chemo-therapy resistant tumour cell; P-glycoprotein; ss.

Homo sapiens.

Key Location/Qualifiers

CDS 425..4267

W08705943-A.

08-OCT-1987.

26-MAR-1987; 87MO-US00758.

01-AUG-1986; 86US-0892575.

28-MAR-1986; 86US-0845610.

(UNIT) UNIV OF ILLINOIS.

Roninson IB, Pastan IH, Gottesman MM;

WPI; 1987-291656/41.

P-PSDB; AAP70452.

DNA for multi-drug resistance in human cells - used to detect

polyptide(s) for diagnosis and therapy

Claim 2(a); Table 5, pp30-39; 61pp; English.

The human multi-drug resistant KB carcinoma cell lines were used as

the source of the mdr1 gene nucleic acid sequences (AA070751). To

obtain cDNA clones of the mdr1 gene (AA070752), poly (A) and RNA was

used. Analysis of the AA sequence presented in (AA070752) indicates

that the mdr1 gene product is likely to be a transmembrane protein.

The presence of transmembrane domains and potential glycosylation

CC sites is consistent with the mdr1 protein being related to the

P-glycoprotein.

Sequence 4669 BP; 1393 A; 896 C; 1126 G; 1254 T; 0 other;

Query Match 78.1%; Score 3341.2; DB 8; Length 4669;

Best Local Similarity 87.6%; Pred. No. 0;

Matches 3758; Conservative 0; Mismatches 488; Indels 44; Gaps 8;

1 GAAGGCGAGGTCCGATCTCTGAAGAGCCGTTAAGGGAGTCCAGAGAAGACTT 60

409 GAGGCGAGGTCCGATCTCTGAAGAGCCGTTAAGGGAGTCCAGAGAAGACTT 60

61 CTGGAATAATGGCAAAAAGTAAAAAATGAGAGAGAAAGAAACCACTGTCTAG 120

469 CTTTCTTAACTGAACAATAAAGTGAAGAGAGAAAGAAACCACTGTCTAG 528

121 CACGTTTCATGTTTGCATTCATCAATGGCTTGATGCTTATATGTTGGTGGGAC 180

529 TGTATTTTCATGTTTGCATTCATCAATGGCTTGATGCTTATATGTTGGTGGGAC 588

181 AATGCTGCGATTCATTCATGAGAGTCCCTCTCATGATGCTGTTTGGAAACAT 240

Db 589 TTTGGCTGCCATCATCCATGAGGCTGAGACTTCCTCATGATCTGTGTGGAGAAAT 648

241 GACAGATAGCTTTGGCAAAATGCGAGAAATTTCAAGAAACAAATTTCCAGTTATATTA 300

649 GACAGATATCTTTGGCAAAATGCGAGAAATTTAGAAATCTGTATGT-----CAAACTAC 702

301 TGAAGTATTTACGAAATATACAAACATTTATCATACACATCTGTGGAGAGAAATACAC 360

703 TAAATGAGAGTATATCAATATACAGAGGTTCTTATCAATATGAGAGAGATACACAG 762

361 GTATGCTATTTATTTACAGTGGATGCTGTGCGCTGTGCTGTGCTTACATCCAGT 420

763 GTATGCTATTTATTTACAGTGGATGCTGTGCGCTGTGCTGTGCTTACATCCAGT 822

421 TTTCAATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 480

823 TTTCAATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 882

481 TGTATATATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 540

883 TGTATATATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 942

541 CCGGCTCAGACAGATGCTGCCAAATCAATGAAGATTTGGGACAAATTTGAGATTT 600

943 CCGGCTCAGACAGATGCTGCCAAATCAATGAAGATTTGGGACAAATTTGAGATTT 1002

601 CTTTCACTCAATGAGCAAAATTTTTCACCGTTTATAGTGGGTTTACAGTGTGGAA 660

1003 CTTTCACTCAATGAGCAAAATTTTTCACCGTTTATAGTGGGTTTACAGTGTGGAA 1062

661 GCTAACCTGCTGATTTTGGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 720

1063 GCTAACCTGCTGATTTTGGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1122

721 AAAGATCTATCTTCTTATTTAGTATTAAGAACTTTGGCTATGCAAAAGCTGAGACGT 780

1123 AAAGATCTATCTTCTTATTTAGTATTAAGAACTTTGGCTATGCAAAAGCTGAGACGT 1182

781 AGCTGAGAGTCTTTAGTATTTAGTATTTAGTATTTAGTATTTAGTATTTAGTATTTAG 840

1183 AGCTGAGAGTCTTTAGTATTTAGTATTTAGTATTTAGTATTTAGTATTTAGTATTTAG 1242

841 ACTGGAAGTATGCAAAATTTTGAAGAGCTTAAAGAGTAAAGAGTAAAGAGTAAAGAGT 900

1243 ACTGGAAGTATGCAAAATTTTGAAGAGCTTAAAGAGTAAAGAGTAAAGAGTAAAGAGT 1302

901 CACGCGCAACATTTCTATTTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 960

1303 TACAGCCAAATTTCTATTTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1362

961 TTTCTGATGAGAGCTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1020

1363 TTTCTGATGAGAGCTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1422

1021 TGTCTGATGAGAGCTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1080

1423 TGTCTGATGAGAGCTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1482

1081 AGCATTTGCAAG 1140

1483 AGCATTTGCAAG 1542

1141 AAGCATTTGCAAG 1200

1543 AAGCATTTGCAAG 1602

1201 ATTCAGAAATGCTTCACTTACCTTCTGAGAAAGAGTAAAGATTTAAAGAGTCT 1260

1603 ATTCAGAAATGCTTCACTTACCTTCTGAGAAAGAGTAAAGATTTAAAGAGTCT 1662

1261 CAACCTGAAGTTCAGAGTGGGACAGAGTGGGCTGTTGGAGAAAGTGGGCTGAGGAA 1320

```
Db 1663 GAACCTGAAAGTGTGACAGAGTGGGACAGCGGTGACCTGTGTGGAAACAGTGGCTGGGAA 1722
Qy 1321 GAGACGAGCCGTGACGTGATGACAGAGGCTCTATGACCCACAGATGGACGTGTAT 1380
Db 1723 GAGACACAGCGTCCAGCTGATGACAGAGGCTCTATGACCCACAGAGGGGATGGCTAGTGT 1782
Qy 1381 TGAATGACAGGACATTAGGACCATTAATTAATTAAGCAATCTGGGAAATTAAGTGGTGT 1440
Db 1783 TGAATGACAGGATATTAGACCATTAATTAATTAAGGTTTCTACGGAAATCAATGGGTGTGT 1842
Qy 1441 GAGTCAGAGGCTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 1500
Db 1843 GAGTCAGAGGACCTGTATGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 1902
Qy 1501 AAATGTCCACCATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1560
Db 1903 AAATGTCCACCATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1962
Qy 1561 CATGAATCTACCTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 1620
Db 1963 CATGAATCTACCTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 2022
Qy 1621 TGAACAGAAACAGAGATGCGCATTTGCTGGGCGCTGTGTGTGTGTGTGTGTGTGTGTGTGT 1680
Db 2023 TGGGACAGAGCAGAGGATGCGCATTTGCTGGGCGCTGTGTGTGTGTGTGTGTGTGTGTGT 2082
Qy 1681 GCTGATGAGGACAGGACGTGACGTGACGTGACGTGACGTGACGTGACGTGACGTGACGTGAC 1740
Db 2083 GCTGATGAGGCGCAGCTGACGTGACGTGACGTGACGTGACGTGACGTGACGTGACGTGAC 2142
Qy 1741 GGAATGAGGCGCAAAAGGCGGAGTACCATTTGATGATGATGATGATGATGATGATGATGATG 1800
Db 2143 GGAATGAGGCGCAAAAGGCGGAGTACCATTTGATGATGATGATGATGATGATGATGATGATG 2202
Qy 1801 TAAATGCCAGTGTGATGCTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 1860
Db 2203 TAAATGCCAGTGTGATGCTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 2262
Qy 1861 TGAATCTCATGAAAGAGAGGCGATTTACTTCAAACTGTGTGTGTGTGTGTGTGTGTGTGTGT 1920
Db 2263 TGAATCTCATGAAAGAGAGGCGATTTACTTCAAACTGTGTGTGTGTGTGTGTGTGTGTGTGT 2322
Qy 1921 TGAATTTGAGTTAGAAATGCGCATGGGATGATGATGATGATGATGATGATGATGATGATG 1980
Db 2323 TGAATTTGAGTTAGAAATGCGCATGGGATGATGATGATGATGATGATGATGATGATGATG 2382
Qy 1981 GTCTCCAAAGATTCAGGCTCCAGTTTAAATTAATTAATTAATTAATTAATTAATTAATTAAT 2040
Db 2383 GTCTCCAAAGATTCAGGCTCCAGTTTAAATTAATTAATTAATTAATTAATTAATTAATTAAT 2442
Qy 2041 TGCACCAACAAGCCAAAGCAGAAAGCTTGTACAAAAGAGAGCACTGATGATGATGATGATG 2100
Db 2443 TGCATTCACAAGCCAAAGCAGAAAGCTTGTACAAAAGAGAGCACTGATGATGATGATGATG 2502
Qy 2101 TCCAGTTTCTCTGAGGATTTCTGAGGCTGAACTCACTGAAATGGGCTTATTTTGTGTGT 2160
Db 2503 TCCAGTTTCTCTTGTGAGGATTTATGAAGCTTAATTTAACTGAAATGGGCTTATTTTGTGTGT 2562
Qy 2161 TGTGATATTTTGTGCTATTTAAACGAGGCTGCAACACAGATTTTAAATTAATTTTGT 2220
Db 2563 TGTGATATTTTGTGCTATTTAAATTAATTAATTAATTAATTAATTAATTAATTAATTTTGT 2622
Qy 2221 AAGGATTAATAGGATCTTTACCCGAGATGAGATCTGTGAACAAACACAGATTAAGTAA 2280
Db 2623 AAGGATTAATAGGATCTTTACCAAGATTAATTAATTAATTAATTAATTAATTAATTAATTA 2682
Qy 2281 CATGTTTCTGATTTGTCTAGCTCTGGAATTAATTTCTTTTAAATTAATTTTCTGCA 2340
Db 2683 CTGTTTCTGATTTGTCTAGCTCTGGAATTAATTTCTTTTAAATTAATTTTCTGCA 2742
Qy 2341 GGGCTTCACATTTGGCAAGGCTGGGAGATCTCACTAAGCGGCTTGCATACATGGTTT 2400
Db 2743 GGGTTTCACATTTGGCAAGGCTGGGAGATCTCACTAAGCGGCTTGCATACATGGTTT 2802

Qy 2401 CAGATCCATGCTGAGACAGATATGACCTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 2460
Db 2803 CCATTCATGCTGACAGATATGATGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 2862
Qy 2461 ATTGAACAACAGGCTGTGCATATGATGTGCGTCAAGTTAAAGGCGCTATAGGTTCCAGGCT 2520
Db 2863 ATTGAACAACAGGCTGTGCATATGATGTGCGTCAAGTTAAAGGCGCTATAGGTTCCAGGCT 2922
Qy 2521 TGTGCTATTTACCAAGATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 2580
Db 2923 TGTGCTATTTACCAAGATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 2982
Qy 2581 TGTGTTGGCAATTAACATCTTTACTGTAGCAATTTGACCAATTTGACCAATTTGACCAATTT 2640
Db 2983 TGTGTTGGCAATTAACATCTTTACTGTAGCAATTTGACCAATTTGACCAATTTGACCAATTT 3042
Qy 2641 TGTGTTGAAATGAAATGTTGCTGTGACACAGCATGAAAGTAAAGTAAAGTAAAGTAAAGTAA 2700
Db 3043 TGTGTTGAAATGAAATGTTGCTGTGACACAGCATGAAAGTAAAGTAAAGTAAAGTAAAGTAA 3102
Qy 2701 TGGGAAGATTTGCTACAGAGCATGAAATCTGCAACCTGTGTGTGTGTGTGTGTGTGTGTGT 2760
Db 3103 TGGGAAGATTTGCTACAGAGCATGAAATCTGCAACCTGTGTGTGTGTGTGTGTGTGTGTGT 3162
Qy 2761 GCAGAAATTTGAATATCATGTATGACAGAGTTTGCAGATACCATACAGAAACCTTTTGA 2820
Db 3163 GCAGAAATTTGAATATCATGTATGACAGAGTTTGCAGATACCATACAGAAACCTTTTGA 3222
Qy 2821 GAAAGACACATCTTGGGCGCTCATTTCTATCATCCAGGCAATGATTTTCTCTA 2880
Db 3223 GAAAGACACATCTTGGGCAATTTCTCTATCATCCAGGCAATGATTTTCTCTA 3282
Qy 2881 TGTGTGCTGTTCGCGTTTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 2940
Db 3283 TGTGTGCTGTTCGCGTTTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 3342
Qy 2941 TGTGTCTTTTGTATTTCTCAGCTATTTGTCTTTTGTGTGTGTGTGTGTGTGTGTGTGTGT 3000
Db 3343 TGTGTCTTTTGTATTTCTCAGCTATTTGTCTTTTGTGTGTGTGTGTGTGTGTGTGTGTGT 3402
Qy 3001 ATTGTGCTGTGATGCAAGGCAAGATATGACAGGCGCCGATGATGATGATGATGATGATG 3060
Db 3403 ATTGTGCTGTGATGCAAGGCAAGATATGACAGGCGCCGATGATGATGATGATGATGATG 3462
Qy 3061 AAAAAGCCTCTGATTTGACAGCTACAGCCTCACGCGCTCAAGCCAAATTTGGAAG 3120
Db 3463 AAAAAGCCTCTGATTTGACAGCTACAGCCTCACGCGCTCAAGCCAAATTTGGAAG 3522
Qy 3121 AAATGTGACATTTAATGAGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 3180
Db 3523 AAATGTGACATTTAATGAGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 3582
Qy 3181 CCAAGGCGGCTGAGGCTGAGGCTGAGGCAAGAGGCGGCGGCGGATGATGATGATGATG 3240
Db 3583 TCAAGGAGCTGAGGCTGAGGCTGAGGCAAGAGGCGGCGGCGGATGATGATGATGATGATG 3642
Qy 3241 CTGTGGGAAAGACACATTTTCACTCTAGAGGCTTCTATGACCCCTTGGTGGTTT 3300
Db 3643 CTGTGGGAAAGACACATTTTCACTCTAGAGGCTTCTATGACCCCTTGGTGGTTT 3360
Qy 3301 AGTGTCAATTTAATGAGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 3360
Db 3703 AGTGTCAATTTAATGAGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 3762
Qy 3361 GGGCATGCTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 3420
Db 3763 GGGCATGCTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 3822
Qy 3421 TGAAGACAAACAGCGGCTGATACATGATGATGATGATGATGATGATGATGATGATGATG 3480
Db 3823 TGAAGACAAACAGCGGCTGATACATGATGATGATGATGATGATGATGATGATGATGATG 3882
```


Db 943 CCGACTTAAAGATGATGCTCTTAAGATTATAGACTATTGGTGCAAAATGGAAAGTT 1002
QY 601 CTTTCACATCATAAGACATTTTTCACCGGTTTATAGTGGGTTTACACGTGGTTGAA 660
Db 1003 CTTTCAGTCAATAGCAATTTTTCACCTGGTTTATAGTAGAGTTTACACGTGGTTGAA 1062
QY 661 GCTAACCTTGGATTGGCCATCAGCCTGTTCTTGAGCTTTGACGTGGCATGGG 720
Db 1063 GCTAACCTTGGATTGGCCATCAGCCTGTTCTTGAGCTTTGACGTGGCATGGG 1122
QY 721 AAGATTAATCTTCTTACTTACTGATTAAGACTCTTGCCCTATGCAAAAGCTGAGCAGT 780
Db 1123 AAGATTAATCTTCTTACTTACTGATTAAGACTCTTGAGGATGCAAAAGCTGAGCAGT 1182
QY 781 AGCTGAAGAGTCTTAGCAGCAATCAGAACTGATTTGCTTTGGAGGCAAAAGAAAGA 840
Db 1183 AGCTGAAGAGTCTTAGCAGCAATCAGAACTGATTTGCTTTGGAGGCAAAAGAAAGA 1242
QY 841 ACTGAAAGATCAACAAATTTAGAAAGCTAAAGGATTTGGGATTAAGAAAGCTAT 900
Db 1243 ACTGAAAGATCAACAAATTTAGAAAGCTAAAGGATTTGGGATTAAGAAAGCTAT 1302
QY 901 CACGGCAACATTTCTATTTGTCGCGCTTCTTATGATCTATGCAATATGCTGCGC 960
Db 1303 TACAGCAATTTCTTATAGGTGCTGCTTCTGCTGATCTATGCTGCTGCGC 1362
QY 961 TTTCTGATGGACCTCTGCTGCTCTCCTCACTGCAATATTTCTATGAGCAACTCTCAC 1020
Db 1363 CTCTGATGGACCTCTGCTGCTCTCCTCACTGCAATATTTCTATGAGCAACTCTCAC 1422
QY 1021 TGTCTCTTTCTGATTAATTTGGGCTTTTATGATTTAGTGAAGCAATCCCAACATTTGA 1080
Db 1423 TGTATTTCTTTCTGATTAATTTGGGCTTTTATGATTTAGTGAAGCAATTTGA 1482
QY 1081 AGCATTTGCAACCCAGAGAGAGAGCTTATGAATCTTCAAGATTAATGCAATTAAC 1140
Db 1483 AGCATTTGCAACCCAGAGAGAGAGCTTATGAATCTTCAAGATTAATGCAATTAAC 1542
QY 1141 AAGCATTTGCAACCCAGAGAGAGCTTATGAATCTTCAAGATTAATGCAATTAAC 1200
Db 1543 AAGCATTTGCAACCCAGAGAGAGCTTATGAATCTTCAAGATTAATGCAATTAAC 1602
QY 1201 ATTCAAAATGTTCACTTCACTTACCTTCTGCAAAAGAGTATGATCTTAAAGGCTT 1260
Db 1603 ATTCAAAATGTTCACTTCACTTACCTTCTGCAAAAGAGTATGATCTTAAAGGCTT 1662
QY 1261 CAACCTGAAGGTTCAAGTGGGACAGAGTGGGCTGTTGGGAAACAGTGGCTGGGAA 1320
Db 1663 CAACCTGAAGGTTCAAGTGGGACAGAGTGGGCTGTTGGGAAACAGTGGCTGGGAA 1722
QY 1321 GAGCAGAGCCGTGAGTGTGATGAGAGGCTCTATGACCCACAGATGGCATGGCTGTAT 1380
Db 1723 GAGCAGAGCCGTGAGTGTGATGAGAGGCTCTATGACCCACAGATGGCATGGCTGTAT 1782
QY 1381 TGAATGAGAGCATTTAGACCAATTAATGTAAGGCACTCTGGGAAATTTACTGGTGGT 1440
Db 1783 TGAATGAGAGCATTTAGACCAATTAATGTAAGGCACTCTGGGAAATTTACTGGTGGT 1842
QY 1441 GAGTCAGAGCCGTGTTGTTGTTGCCACACAGATGGTGAACATTTGCTATGGCCGCA 1500
Db 1843 GAGTCAGAGCCGTGTTGTTGTTGCCACACAGATGGTGAACATTTGCTATGGCCGCA 1902
QY 1501 AATGTCACATGATGATGATGAGAAAGCTTTAAGAAAGCATTTGGCTATGATTTAT 1560
Db 1903 AATGTCACATGATGATGATGAGAAAGCTTTAAGAAAGCATTTGGCTATGATTTAT 1962
QY 1561 CATGAAGTACTTAATTAATTTGACACTCTGGTTGGAGAGAGGGCCGCTGAGTGG 1620
Db 1963 CATGAAGTACTTAATTAATTTGACACTCTGGTTGGAGAGAGGGCCGCTGAGTGG 2022
QY 1621 TGGAGCAAGAGAAATCGCATTTGCTGGGCTGTTGCCAACCCAGATTTCTTCT 1680
Db 1681 TGGAGCAAGAGAAATCGCATTTGCTGGGCTGTTGCCAACCCAGATTTCTTCT 2082
QY 1681 GCTGATGAGAGCAACGTGAGCTCTTGACACTGAAGTGAAGCATGTTGAGTGGCCCT 1740
Db 2083 GCTGATGAGAGCAACGTGAGCTCTTGACACTGAAGTGAAGCATGTTGAGTGGCCCT 2142
QY 1741 GGATTAAGGCAAGAAAGGCGACTATCATTTGATGAGTCACTGTTGTTACAGTTG 1800
Db 2143 GGATTAAGGCAAGAAAGGCGACTATCATTTGATGAGTCACTGTTGTTACAGTTG 2202
QY 1801 TTAATGCGATGCTATGCTGCTTTTATGATGATGAGTCACTTTGGAGAAAGAAATCTGA 1860
Db 2203 TTAATGCGATGCTATGCTGCTTTTATGATGATGAGTCACTTTGGAGAAAGAAATCTGA 2262
QY 1861 TGAACATGAAGAGAGAGGCAATTTACTTCAAACTTTGCAACATGACAGAACAGGAAA 1920
Db 2263 TGAACATGAAGAGAGAGGCAATTTACTTCAAACTTTGCAACATGACAGAACAGGAAA 2322
QY 1921 TGAATTTGATTAAGAAATGTCACATGGTGAATCCAAAGTGAATGCTTGGCAAT 1980
Db 2323 TGAATTTGATTAAGAAATGTCACATGGTGAATCCAAAGTGAATGCTTGGCAAT 2382
QY 1981 GTCTCCAAAGATTCAGGCTCAGCTTATTAAGAAAGATCACTCCAGAGATTAACA 2040
Db 2383 GTCTCCAAAGATTCAGGCTCAGCTTATTAAGAAAGATCACTCCAGAGATTAACA 2442
QY 2041 TGCACACAGAGCCCAAGACAGAAAGCTTGGTACAAAGAGACTTGAATGGAATGACC 2100
Db 2443 TGCACACAGAGCCCAAGACAGAAAGCTTGGTACAAAGAGACTTGAATGGAATGACC 2502
QY 2101 TCCAGTTCTCTTGGAGAGATTCGAACTGAACTCACTGAAATGGCTTTTGTGGT 2160
Db 2503 TCCAGTTCTCTTGGAGAGATTCGAACTGAACTCACTGAAATGGCTTTTGTGGT 2562
QY 2161 TGGTATATTTTGTGCTATTAATTAAGAGAGGCTGCACACAGCATTTTCAATATTTTC 2220
Db 2563 TGGTATATTTTGTGCTATTAATTAAGAGAGGCTGCACACAGCATTTTCAATATTTTC 2622
QY 2221 AAGGATTAAGGATCTTTTACCCAGATGAGATCCTGAAACAAACAAACAGATTAAGTA 2280
Db 2623 AAGGATTAAGGATCTTTTACCCAGATGAGATCCTGAAACAAACAAACAGATTAAGTA 2682
QY 2281 CATGTTTCTGATTTGTTTCTGAGTCTGAGATTAATTTCTTTTATTAATTTTCCGCA 2340
Db 2683 CATGTTTCTGATTTGTTTCTGAGTCTGAGATTAATTTCTTTTATTAATTTTCCGCA 2742
QY 2341 GGGCTTACATTTGGCAAGCTGGGAGATCTCACTAAGCGCTTGGATCATGATGTTTT 2400
Db 2743 GGGCTTACATTTGGCAAGCTGGGAGATCTCACTAAGCGCTTGGATCATGATGTTTT 2802
QY 2401 CAGATCATGCTGAGACAGATGTCAGAGTGTGATGACCCATAAAACACACATGGAGC 2460
Db 2803 CAGATCATGCTGAGACAGATGTCAGAGTGTGATGACCCATAAAACACACATGGAGC 2862
QY 2461 ATTGAACAACAGGCTTGGCAATGATGGGCTCAAGTAAAGAGGCTTAAGGTTCCAGGCT 2520
Db 2863 ATTGAACAACAGGCTTGGCAATGATGGGCTCAAGTAAAGAGGCTTAAGGTTCCAGGCT 2922
QY 2521 TGTCTCATTTACCCAGAAATATAGCAAACTTTGGAGACGAGCATTAATATATCTTAATCTA 2580
Db 2923 TGTCTCATTTACCCAGAAATATAGCAAACTTTGGAGACGAGCATTAATATATCTTAATCTA 2982
QY 2581 TGGTTGGCAATTAACACTTTACTTACCAATTTGACCAATTTGCAATTAAGCAGAGT 2640
Db 2983 TGGTTGGCAATTAACACTTTACTTACCAATTTGACCAATTTGCAATTAAGCAGAGT 3042
QY 2641 TGTGTAATGAAGATTTGTTGTTGTCGACAGACCTGAAAGTAAAGAAAGCTAGAAAGAGC 2700
Db 3043 TGTGTAATGAAGATTTGTTGTTGTCGACAGACCTGAAAGTAAAGAAAGCTAGAAAGAGC 3102
QY 2701 TGGAGAGATTTGCTACAGAGCCTGAAACCTTCCGAATCTGTTGTTCTTTGACTGGGA 2760
Db 3103 TGGAGAGATTTGCTACAGAGCCTGAAACCTTCCGAATCTGTTGTTCTTTGACTGGGA 3162

Accession	Gene	Position	Sequence	Length
3641	CCAGGCTGAGGAAAGCGCTAGTGAACCTGTGGCCATATGAGCTGTTAAATATTTTAACT	3900		
4243	CCAGGCTGAGGAAAGCGGCGCACTGAACCTGACTGTATGAGATGTTAAATCTTTAAAT	4302		
3901	ATTGTGTTAAACATGCACTTTAAATCAAAAGTTAAAGTGAGCACTTACTGGAAAACT	3960		
4303	ATT--TGTTAGATATGACATTTATTCAAAGTTAA-----AGCAACACTTACAGAAAT	4355		
3961	ATGTAGAAACACCGTTTAACTTCTCTGTCGACACGTAGATCATTCACCAAGTTGAG	4020		
4356	ATGAAGAGGATCTGTTTAACATTTCTCAGTCAGATTCAGAGCTTC-----4403	4403		
4021	AGCTTCAGATTTTAAATTAAGAGAACCA--AAGAAACATTAATGANGAATAAAT	4078		
4404	-----AGAGACTTCGTATTAAGGAACAGAGTGAGACATCATCAAGTGAGAGAAAT	4458		
4079	ACTGCTGTTATTCGATTAATAAATTATAGAGTAATTCAAAGTACATTTTGTAAATAT	4138		
4459	CATGATTTTAACTGCACTTAATTAATTTATATACGATATTAAGATATTTTAAAGATAA	4518		
4139	TGTATAATTTTGTATATATTT-----ATTGTACTTACGTCTCTGTAAGAT	4190		
4519	ATGTGTAATTTGTTATATATTTTCCACTTGGACTGTACAGTCCCTTGTCTAAAGAT	4578		
4191	TATAGAAAGGTAAAGAGTACTG--AATGTTGAATAAAGTGCTAGCTATATTAACCTAA	4249		
4579	TATAGAAAGGTAAAGAGTATGAAGATGTTGCATTAAGTG--TCTATATTAACCTAA	4635		
4250	ACTTTATATCAAAAAAAAAAAAAAAAAAAAA 4279			
4636	ACTTTCATCGAAAAAAAAAAAAAAAAAAAAA 4665			

2101 TCCAGTTTCCTCTGAGAGATTCGAAAGCTGAACTCACTGAATGAGCCCTTATTTGCTG 2160
2216 TCCAGTTTCCTCTGAGAGATTCGAAAGCTGAACTCACTGAATGAGCCCTTATTTGCTG 2275
2161 TGGTATATTTTGGCTATATTAACGAGAGCCCTGCAACGACATTTTCATATATTTTC 2220
2276 TGGTATATTTTGGCTATATTAACGAGAGCCCTGCAACGACATTTTCATATATTTTC 2335
2221 AAGGATTAATGAGATCTTACCCGAGATCCGTAAGCAACAAACGACAGAAATAGTAA 2280
2336 AAGGATTAATGAGATCTTACCCGAGATCCGTAAGCAACAAACGACAGAAATAGTAA 2395
2281 CATGTTTCTGATGTTGTTCTAGCTCTGGAATTAATTTCTTTATTTACATTTTTCCTCA 2340
2396 CTGTTTTCATATTTGTTCTAGCTCTGGAATTAATTTCTTTATTTACATTTTTCCTCA 2455
2341 GGGCTTCACATTTGGCAAGCTGGGGAGATCCTCACTAAGGGGCTTCGATCATGTTTTT 2400
2456 GGGTTTCACATTTGGCAAGCTGGGGAGATCCTCACTAAGGGGCTTCGATCATGTTTTT 2515
2401 CAGATTCATGCTGAGACAGAGATGTCAGCTGTTGATGACCTTAATAAACACACCTGGAGC 2460
2516 CCGATTCATGCTGAGACAGAGATGTCAGCTGTTGATGACCTTAATAAACACACCTGGAGC 2575
2461 ATTGACAAACAGGCTTGCATGATGAGCAATGTCGCTCAAGTTAAAGGGGCTATAGTTCCAGCT 2520
2576 ATTGACAAACAGGCTTGCATGATGAGCAATGTCGCTCAAGTTAAAGGGGCTATAGTTCCAGCT 2635
2521 TGCCTGATATCCAGAAATATAGCAAAATCTGGGACAGGATATATATTCCTTATCTA 2580
2636 TGCCTGATATCCAGAAATATAGCAAAATCTGGGACAGGATATATATTCCTTATCTA 2695
2581 TGGTGGCAATTAACATTTTACTTATAGCAATTTGACCATTCATTCGAATAGCAGAGT 2640
2696 TGGTGGCAATTAACATTTTACTTATAGCAATTTGACCATTCATTCGAATAGCAGAGT 2755
2641 TGGTGAATGAATAATGTTGCTGACAAAGCACTGAAGATTAAGAGGCTAGAGAGC 2700
2756 TGGTGAATGAATAATGTTGCTGACAAAGCACTGAAGATTAAGAGGCTAGAGAGC 2815
2701 TGGGAAGATTTGCTACAGAAAGCATGAAACCTCCGAAGTGTCTTTTGACTGGGA 2760
2816 TGGGAAGATTTGCTACAGAAAGCATGAAACCTCCGAAGTGTCTTTTGACTGGGA 2875
2761 GCAGAGTTTGAATACATATGACAGAGTTTGCAGTACCATACAGAAACCTTTGAG 2820
2876 GCAGAGTTTGAATACATATGACAGAGTTTGCAGTACCATACAGAAACCTTTGAG 2935
2821 GAAAGCACATCTTGGGGGCTCATTTTCTATCACCCAGGCAATGATATTTTCTTA 2880
2936 GAAAGCACATCTTGGGAATTCATTTTCTATCACCCAGGCAATGATATTTTCTTA 2995
2881 TGTGCTGTTTCCGGTTTGGTGGCTACTTGTGGCAATGATGATGATTCAGAGA 2940
2996 TGTGCTGTTTCCGGTTTGGAGCTTGTGGGACATTAACATGAGAGCTTTGAGA 3055
2941 TGTCTTTTGGTATTCACATCTATGCTTGGTGCATGAGCAATGGGCAAGTTC 3000
3056 TGTCTGTTAGTATTTCAAGCTGTGTGCTTTGGTGCATGAGGCGAGGCAAGTTC 3115
3001 ATTTGCTCTGACTATGCAAGGCAAGATATGACAGCCGACATCATCATGATCATTTGA 3060
3116 ATTTGCTCTGACTATGCAAGGCAAGATATGACAGCCGACATCATCATGATCATTTGA 3175
3061 AAAAGCCCTCTGATGACAGCTACAGCCCTCAAGGCTCAAGCCAAATACGTTGAAAG 3120
3176 AAAAGCCCTCTGATGACAGCTACAGCCGACAGGAGGCTTAATCGAACAATGGAAG 3235
3121 AATATGACATTTAATGAGTCTGTTCAACTATCCACTGACAGACATCCCGTGT 3180
3236 AATATGACATTTGTTGAAGTTGATTCATATCCACCAGACCGACATCCAGTGT 3295
3181 CCAGGGGCTGAGCTGAGGTGAAGAGGCGCAAGCGCTGCCCTGTAGTAGCAGTGG 3240

3296 TCAGGAGCTGAGCTGAGAGTGAAGAGGCCAGAGCGTGGCTGGTGGGAGCAGTGG 3355
3241 CTGTTGGAGAGACAGATGTTTACCTCTAGAGCGCTTATATACCCCTTGGCTGTTTC 3300
3356 CTGTTGGAGAGACAGATGTTTACCTCTAGAGCGGTTTACAGACCCCTTGGAGGAA 3415
3301 AGTGTATTTGATGCAAGAGATTAAGCACTGAATGCTCAGAGGCTCCGACACCT 3360
3416 AGTGTCTCTGATGCAAGAGATTAAGCACTGAATGCTCAGAGGCTCCGACACCT 3475
3361 GGGATGCTGCTCAGAGAGCCCATCTGTTTGAATGCTGACAGATTTGCCGGAACATTTGCTA 3420
3476 GGGATGCTGCTCAGAGAGCCCATCTGTTTGAATGCTGACAGATTTGCCGGAACATTTGCTA 3535
3421 TGGAGACACAGCCGGTCTGATCATATGAAGATTAATGACAGGAGCCAA 3480
3536 TGGAGACACAGCCGGTCTGATCATATGAAGATTAATGACAGGAGCCAA 3595
3481 CATACACATTCATGAGACACTCCTGAGAAATACAAACCCAGATGAGAGCAAAAG 3540
3596 CATACATTCATGAGACACTCCTGAGAAATACAAACCCAGATGAGAGCAAAAG 3655
3541 AACCCAGCTCTGCTGAGCAAAAGCCGATGAGCAATGAGCTGAGCTGTTAGACA 3600
3656 AACCCAGCTCTGCTGAGCAAAAGCCGATGAGCAATGAGCTGAGCTGTTAGACA 3715
3601 GCGTATATTTGCTTGTGATGAGAGTACATCAGCTCTGATACGAAAGTGAAGGT 3660
3716 GCGTATATTTGCTTGTGATGAGAGTACATCAGCTCTGATACGAAAGTGAAGGT 3775
3661 TGTCCAAGAGCCCTGAGCAAAAGCCGATGAGCAATGAGCTGAGCTGTTAGACA 3720
3776 TGTCCAAGAGCCCTGAGCAAAAGCCGATGAGCAATGAGCTGAGCTGTTAGACA 3835
3721 CTGTTCCACATTCAGAAATGCAATTAATAGTGTGTTTCAAGATGCAAGAGT 3780
3836 CTGTTCCACATTCAGAAATGCAATTAATAGTGTGTTTCAAGATGCAAGAGT 3895
3781 GCATGACACATCAACAGCTCTGAGCCCAAGAGGCACTATATTTTCCATGCTAGT 3840
3896 GCATGACACATCAACAGCTCTGAGCCCAAGAGGCACTATATTTTCCATGCTAGT 3955
3841 CCAGGCTGAGCAAGGCTGATGAACTGTGGCATATGAGCTGTTAAATTTTAAAT 3900
3956 CCAGGCTGAGCAAGGCTGATGAACTGTGGCATATGAGCTGTTAAATTTTAAAT 4015
3901 ATTTGCTTAAACATGCAATTAACAAAGTTAAAGGTCAGCACTTACTGCAAAACT 3960
4016 ATTTGCTTAAACATGCAATTAACAAAGTTAAAGGTCAGCACTTACTGCAAAACT 4068
3961 ATTTGCTTAAACATGCAATTAACAAAGTTAAAGGTCAGCACTTACTGCAAAACT 4020
4068 ATTTGCTTAAACATGCAATTAACAAAGTTAAAGGTCAGCACTTACTGCAAAACT 4116
4069 ATTTGCTTAAACATGCAATTAACAAAGTTAAAGGTCAGCACTTACTGCAAAACT 4171
4021 ATTTGCTTAAACATGCAATTAACAAAGTTAAAGGTCAGCACTTACTGCAAAACT 4078
4117 ATTTGCTTAAACATGCAATTAACAAAGTTAAAGGTCAGCACTTACTGCAAAACT 4171
4079 ATTTGCTTAAACATGCAATTAACAAAGTTAAAGGTCAGCACTTACTGCAAAACT 4138
4172 ATTTGCTTAAACATGCAATTAACAAAGTTAAAGGTCAGCACTTACTGCAAAACT 4231
4139 ATTTGCTTAAACATGCAATTAACAAAGTTAAAGGTCAGCACTTACTGCAAAACT 4190
4232 ATTTGCTTAAACATGCAATTAACAAAGTTAAAGGTCAGCACTTACTGCAAAACT 4291
4191 ATTTGCTTAAACATGCAATTAACAAAGTTAAAGGTCAGCACTTACTGCAAAACT 4249
4292 ATTTGCTTAAACATGCAATTAACAAAGTTAAAGGTCAGCACTTACTGCAAAACT 4348
4250 ATTTGCTTAAACATGCAATTAACAAAGTTAAAGGTCAGCACTTACTGCAAAACT 4279

Db 4349 ACTTCATGTGACTGGAAAAA 4378

RESULT 9

AAV32645 ID AAV32645 standard; cDNA: 4669 BP.

AC AAV32645;

XX 23-SEP-1998 (first entry)

XX Human P glycoprotein (Pgp) cDNA.

XX Human P glycoprotein; Pgp; multi-drug resistance; cancer;

XX UIC2 monoclonal antibody; mAb; cytotoxic; transmembrane efflux pump; ss.

XX Homo sapiens.

XX Key Location/Qualifiers

XX 5'UTR 1..424

XX CDS /tag= a

XX /tag= b

XX 3'UTR 4268..4669

XX /tag= c

XX MO9821325-A1.

XX 22-MAY-1998.

XX 17-NOV-1997; 97WO-US21214.

XX 15-NOV-1996; 96US-0752447.

XX (INGE-) INGENEX INC.

XX (UNIT) UNIT ILLINOIS FOUND.

XX Mechtner E, Roninson IB;

XX WPI: 1998-297930/26.

XX P-PSDB: AAM48997.

XX Immunological reagent specific for P-glyco:protein - useful for

XX detecting multi-drug resistant cancer, isolating haematopoietic

XX cells and selective cell killing

XX Claim 1; Fig 1A; 89pp; English.

XX The present sequence represents the human P glycoprotein (Pgp)

XX cDNA which encodes the wild-type Pgp protein. Pgp is a transmembrane

XX efflux pump protein involved in multi-drug resistance of cancer cells.

XX The invention provides methods for developing and using immunological

XX reagents specific for certain mutant forms of Pgp and wild-type Pgp in

XX a conformation associated with substrate binding or in the presence

XX of ATP depleting agents. An example of the immunological reagent

XX is the UIC2 monoclonal antibody (mAb). mAb UIC2 specifically binds

XX to Pgp in a particular biochemical conformation and is capable of

XX inhibiting drug efflux from Pgp-expressing cells. The immunological

XX mammalian cells, including low level expression, particularly in

XX cells to diagnose multi-drug resistance. The invention claims that

XX these immunological reagents are more specific than known reagents for

XX detecting Pgp and they also eliminate the need for costly and laborious

XX screening of Pgp inhibitors by cytotoxicity or dye exclusion methods.

XX Sequence 4669 BP; 1393 A; 894 C; 1130 G; 1252 T; 0 other;

Query Match 77.9%; Score 333.2; DB 19; Length 4669;

Best Local Similarity 87.5%; Pred. No. 0; Mismatches 493; Indels 44; Gaps 8;

1 GGAGCGGAGGTGCGGATGATCTGAAGGAGCGCGTAAGGAGGTGACAGAAAGAACTT 60

Db 409 GGAGCGGAGGTGCGGATGATCTTGAAGGAGCGCGTAAGGAGGTGACAGAAAGAA 468

Qy 61 CTGGAATAATGGGCAAAAAAGTAAAAAATGAGAGAAAGAAAGAAAGCAATCTGAC 120

Db 469 CTTTAAAACTGAACATTAAGTGAAGAAAGTGAAGAAAGAAAGCAATCTGAC 528

Qy 121 CACGTTGCAATGTTGGCTATCAAAATGCTGATAGTGTATATGTTGGTGGGAC 180

Db 529 TGTATTTCAATGTTTGGCTATCAAAATGCTGATAGTGTATATGTTGGTGGGAC 588

Qy 181 AATGCTGCATCATCCATGAGCTGACCTCCCTCATGATGCTGTTTGGAAACAT 240

Db 589 TTGGCTGCATCATCCATGAGCTGACCTCCCTCATGATGCTGTTTGGAAACAT 648

Qy 241 GACAGATAGCTTGGCAAAAGGAGAAATTTCAAGAAACAAACTTTCCAGTTATATTA 300

Db 649 GACAGATATCTTGGCAAAAGGAGAAATTTAGAGATCTGATGT-----CAACATCAC 702

Qy 301 TGAAGTATTACGACATACATACATTTATCATCAACCATCGGAGGAGAAAGACAC 360

Db 703 TAATAGAGTATATCATATGATATAGAGGTTCTTATGATCTGAGAGAGACACG 762

Qy 361 GTATGCTATATTATACAGTGGATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 420

Db 763 GTATGCTATATTATACAGTGGATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 822

Qy 421 TTCATTCGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 480

Db 823 TTCATTCGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 882

Qy 481 TGCATATGTCGACAGAGAGATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 540

Db 883 TGCATATGTCGACAGAGAGATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 942

Qy 541 CCGGCTCACAGAGATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 600

Db 943 CCGGCTCACAGAGATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1002

Qy 601 CTTTACATCAATAGCAATTTTACCGGTTTATATGATGCTGCTGCTGCTGCTGCTGCT 660

Db 1003 CTTTACATCAATAGCAATTTTACCGGTTTATATGATGCTGCTGCTGCTGCTGCTGCT 1062

Qy 661 GCTAACCTTGTGATTTTGGGATCAGCCCTGTTCTTGGATTCAGCCGCTGCTGCTGCT 720

Db 1063 GCTAACCTTGTGATTTTGGGATCAGCCCTGTTCTTGGATTCAGCCGCTGCTGCTGCT 1122

Qy 721 AAGATATCTATCTTCAATTTTACGATAAAGAACTTGGGCTATGCAAGCTGAGCAGT 780

Db 1123 AAGATATCTATCTTCAATTTTACGATAAAGAACTTGGGCTATGCAAGCTGAGCAGT 1182

Qy 781 AGCGAAGAGTCTTACAGCAATCAAGCTGATGCTGCTTGGAGGACAAAGAAAGA 840

Db 1183 AGCGAAGAGTCTTACAGCAATCAAGCTGATGCTGCTTGGAGGACAAAGAAAGA 1242

Qy 841 ACTTGAAGGTACAAACAAATTTAGAGAGTAAAGAAATTTGGATTTGGAATGCTAT 900

Db 1243 ACTTGAAGGTACAAACAAATTTAGAGAGTAAAGAAATTTGGATTTGGAATGCTAT 960

Qy 901 CAGGCGCAATTTTCAATTTGCGCTTCTTATGATCTATGATCTATGATCTATGATCT 960

Db 1303 TACAGCAATTTTCAATTTGCGCTTCTTATGATCTATGATCTATGATCTATGATCT 1362

Qy 961 TTTCTGTATGGGACCTCTTGTGCTCTCCGAGTAATTTCTATGATCTATGATCTATG 1020

Db 1363 CTTCTGTATGGGACCTCTTGTGCTCTCCGAGTAATTTCTATGATCTATGATCTATG 1422

Qy 1021 TGTCTTCTTCTTGTATTAATTTGAGGCTTTATGATGAGACGATCCCAAGCATTTGA 1080

Db 1423 TGTATCTTCTTGTATTAATTTGAGGCTTTATGATGAGACGATCCCAAGCATTTGA 1482

Qy 1081 AGCATTTGCAAGCGAAGGAGGAGCTTATGAATCTTCAAGATTAATGACATAAACC 1140

Db 1483 AGCATTTGCAAAATGCAAGAGAGACGCTTATGAAATCTTCAAGATTAATTTGATTAATTAAGCC 1542
Oy 1141 AAGCATTTGACAGTATTTGAGAGATGACATTAACCCAGATTAATTTAAGGAAATTTGGA 1200
Db 1543 AAGTATTGACAGCTATTTCAGAGAGTGGGACAAACCCAGATTAATTTAAGGAAATTTGGA 1602
Oy 1201 ATTCAAAAATGTTCACTTCAGTTTACCTTCTCTGAAAAGAAAGTTAGATCTTAAGGGCT 1260
Db 1603 ATTCAAGAAATGTTCACTTCAGTTTACCTTCTGAAAAGAAAGTTAGATCTTGAAGGGCT 1662
Oy 1261 CAACCTGAAGCTTCAGAGTGGGACAGACATGGCGCTGTTGGGAAACAGTGGCTCGGGAA 1320
Db 1663 GAACCTGAAGCTTCAGAGTGGGACAGACGGGCGCTGTTGGAAACAGTGGCTCGGGAA 1722
Oy 1321 GAGCAGACCGGTGACGTATGACAGAGCTCTATGACCCCAAGATGGATGGTCTGAT 1380
Db 1723 GAGCAGAAACGTCCAGCTGATGACAGAGGCTCTATGACCCCAAGAGGATGGTCTGAT 1782
Oy 1381 TGATGACAGGACATTTAGGACCAATTAATGTAAGCATCTTTCGGGAAATTAAGTGGTGT 1440
Db 1783 TGATGACAGGATATTAGACCATTAATGTAAGTTTCTAAGGAAATCATTTGGTGTGT 1842
Oy 1441 GAGTCAGAGGCTGTGTTGTTTGGCCACACAGATAGCTGAAAACATTCGCTATGGCCGGA 1500
Db 1843 GAGTCAGGAACCTGTATTTGTTGGCCACACAGATAGCTGAAAACATTCGCTATGGCCGTA 1902
Oy 1501 AAATGTCACCTGATGATGATTTGAAAAGCGTTTAAAGGAAGCAATTCCTATGATTTTAT 1560
Db 1903 AAATGTCACCTGATGATGATTTGAAAAGCGTTTCAAGAGAAAGCCAAATTCCTATGATTTAT 1962
Oy 1561 CATGAACCTACCTAATTAATTTGACACTGTGTTGAGAGAGAGGGGCCCAAGCTGATG 1620
Db 1963 CATGAACCTGCTCATTAATTTGACACCTGTGTTGAGAGAGAGGGGCCCAAGCTGATG 2022
Oy 1621 TGGCAGAAACAGAGAAATCCGATTTGCTCGGGCCCTGGTTGCAACCCCAAGATTTCTCT 1680
Db 2023 TGGGAGAGAGAGAGATGCCATTTGCTGACGTGCTGCTGTTGCAACCCCAAGATTTCTCT 2082
Oy 1681 GCTGGATGAGGCAAGCTGACGTCTGACACTGAAAAGTGAAGCAGTGGTTCAGGTGGCT 1740
Db 2083 GCTGGATGAGGCAAGCTGACGTCTGACACTGAAAAGCAGTGGTTCAGGTGGCTCT 2142
Oy 1741 GGATTAAGCCAGAAAAGCCGAGCTACACATTTGATAGCTCATGCTTGTGTTGATAGTTTCG 1800
Db 2143 GGATTAAGCCAGAAAAGGTCGAGCACTTTGATAGCTCATGCTTGTGTTGATAGTTTCG 2202
Oy 1801 TAATGCCATGTCATTTGCTGTTTGTGATGATGAGTCAATTTGAGAAAAGAAATCATGA 1860
Db 2203 TAATGCCATGTCATTTGCTGTTTGTGATGATGAGTCAATTTGAGAAAAGAAATCATGA 2262
Oy 1861 TGAATCATGAAAAGAGAGGCAATTTACTTCAAACTGTGCACAAATGACAGACAAAGAGAA 1920
Db 2263 TGAATCATGAAAAGAGAGGCAATTTACTTCAAACTGTGCACAAATGACAGAGAGAGAA 2322
Oy 1921 TGAATTCATTTGAAAATCCACTGTGATTCAAAATGAAAAGTGAAGTATGATGCTTGAAT 1980
Db 2323 TGAATTCATTTGAAAATCCACTGTGATTCAAAATGAAAAGTGAAGTATGATGCTTGAAT 2382
Oy 1981 GTCTCCAAAAGATTCAGGTCCAGTATTAAATTAAGAAAGATCAATCGAGAGATATACA 2040
Db 2383 GTCTCCAAAAGATTCAGAGTCCAGTATTAAATTAAGAAAGATCAATCGAGAGAGTCCG 2442
Oy 2041 TGCACCAAGGCCAAGACAGAAAGCTTGTACAAAAGAGAGACTTGATAGATGATACC 2100
Db 2443 TGCACCAAGGCCAAGACAGAAAGCTTGTACAAAAGAGAGCTTGTAGAAAGTATACC 2502
Oy 2101 TCCAGTTTCCTTGGAGAGATTCGAAGCTGAACCTCAACAGATTTGCTATTTTGTGT 2160
Db 2503 TCCAGTTTCCTTGGAGAGATTCGAAGCTGAACCTCAACAGATTTGCTATTTTGTGT 2562
Oy 2161 TGGTATATTTTGTGCTATTAATTAACGGAGGCTGCACACAGATTTTTCATTAATTTTC 2220
Db 2563 TGGTATATTTTGTGCTATTAATTAATTAAGAGGCTGCACACAGATTTTTCATTAATTTTC 2622

Oy 2221 AAGATTTATAGGATCTTTACCCAGATGAGATCTGAAAACAAAACGACAGATATGTA 2280
Db 2623 AAGATTTATAGGGGTTTATTAACAGATTTGATGATCTGAAAACAAAACGACAGATATGTA 2682
Oy 2281 CATGTTTCTGATGATTTCTCTAGCTTGTGAATTAATTTCTTTTATTAATTTTCTTCA 2340
Db 2683 CTGTTTCTGATGATTTCTCTAGCTTGTGAATTAATTTCTTTTATTAATTTTCTTCA 2742
Oy 2341 GGGCTTCACATTTGGCAAGCTGGGAGATCTCTCAATGAGGGGCTGATCATGTTT 2400
Db 2743 GGGCTTCACATTTGGCAAGCTGGGAGATCTCTCAACAGAGGCTCGATCATGTTT 2802
Oy 2401 CAATTCAGTCTGAGACAGATATGTCAGTGTGTTGATGACCTAATAAACCCACTGAGC 2460
Db 2803 CCAATTCAGTCTGAGACAGATATGTTGTTGTTATGACCTAATAAACCCACTGAGC 2862
Oy 2461 ATTGACAAACGCTTGCACATGATCGGCTCAAGTTAAAGGGGCTATAGTTCCAGCT 2520
Db 2863 ATTGACAAACGCTTGCACATGATCGCTCAAGTTAAAGGGGCTATAGTTCCAGGCT 2922
Oy 2521 TGTCTGATTAACCAAGATATGACAAATCTTGGACAGGCAATTAATTAATCTTAATCTA 2580
Db 2923 TGTCTGATTAACCAAGATATGACAAATCTTGGACAGGAATTAATTAATCTTAATCTA 2982
Oy 2581 TGTGTCACATTAACACTTTTACTTTAGCATTTGATACCATCATTTGCAATGACAGAGT 2640
Db 2983 TGTGTCACATTAACACTTTTACTTTAGCATTTGATACCATCATTTGCAATGACAGAGT 3042
Oy 2641 TGTGTAATTAATTAATTTGTTGCTGACAGACACTGAAAAGATTAAGAAAGCTAGAAAGAC 2700
Db 3043 TGTGTAATTAATTAATTTGTTGCTGACAGACACTGAAAAGATTAAGAAAGCTAGAAAGTGC 3102
Oy 2701 TGGGAAGATTTCTACAGAAAGCCATCGAAAACCTTCGAATCTGTTGTTTGTGATCTGGGA 2760
Db 3103 TGGGAAGATTTCTACAGAAAGCCATCGAAAACCTTCGAATCTGTTGTTTGTGATCTGGGA 3162
Oy 2761 GCAGAGTTTGAATTAATGATGATGACAGAGTTTGCATGACATCCATAGCAATCTTTGAG 2820
Db 3163 GCAGAGTTTGAATTAATGATGATGACAGAGTTTGCATGACATCCATAGCAATCTTTGAG 3222
Oy 2821 GAAAGCAGACATCTTCGGGGTCTCATTTTCTATCACCCAGGCAATGATTTTCCCTA 2880
Db 3223 GAAAGCAGACATCTTTGGAATTAATTTCTTCACCCAGGCAATGATTTTCCCTA 3282
Oy 2881 TGTGCTGTTTCCGCTTGTGCTTGTGCTTGTGCTTGTGCTTGTGCTTGTGCTTGTGCT 2940
Db 3283 TGTGCTGTTTCCGCTTGTGCTTGTGCTTGTGCTTGTGCTTGTGCTTGTGCTTGTGCT 3342
Oy 2941 TGTGCTTGTGATTTCTCAGTATTTGCTTGTGCTTGTGCTTGTGCTTGTGCTTGTGCT 3000
Db 3343 TGTGCTTGTGATTTTCTCAGTATTTGCTTGTGCTTGTGCTTGTGCTTGTGCTTGTGCT 3402
Oy 3001 ATTTGCTCTGACTATGCAAAAGCCAAAGATATACAGACCCCAAGCATCATGATCATTTGA 3060
Db 3403 ATTTGCTCTGACTATGCAAAAGCCAAATATACAGACCCCAAGCATCATGATCATTTGA 3462
Oy 3061 AAAAAGCCCTGATTTGACAGCTACAGCCCTCAGGCCCTCAAGCCAAATACGTTGGAAG 3120
Db 3463 AAAAAGCCCTTATTTGACAGCTACAGCAAGGAGCCCTTAAGCCGAACATGTTGGAAG 3522
Oy 3121 AAATGTCACATTTAATGAGTGTGTTCAACTATCCCACTGACACAGACATCCCGTGTCT 3180
Db 3523 AAATGTCACATTTGGAATTTGATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 3582
Oy 3181 CCAGGGCTGAGGCTGAGAGTGAAGAAGGGCCAGACGTGGCCCTGTAGTATGACATGG 3240
Db 3583 TCAGGGACTGAGGCTGAGAGTGAAGAAGGGCCAGAGGTGGCTGTAGTATGACATGG 3642
Oy 3241 CTGTGGAGAGACAGATTTTCAAGCTCTAGAGCGCTTCTATGACCCCTTGGCTGGTTC 3300
Db 3643 CTGTGGAGAGACAGATTTTCAAGCTCTAGAGCGCTTCTATGACCCCTTGGCTGGAGAA 3702

```

OY 3301 AGTGTAAATGATGCAAAAGATPAAGACCTGAAATGTCAGTGGCTCCAGCACACT 3360
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 3703 AGTGTCTGCTGTGATGCAAAAGAAATAAGCAGTAATGTTCAAGTGGCTCCAGCACACT 3762
OY 3361 GGGCATGCTGTGTCTCAGAGAGCCATCTGTTTGACTGACATTTGCCGGAACATTTGCTA 3420
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 3763 GGGCATGCTGTGTCTCAGAGAGCCATCTGTTTGACTGACATTTGCCGGAACATTTGCTA 3822
OY 3421 TGGAGAACACAGCCGGGTGTATCAGATGAAGAAATTAATTCAGAGCCAGGAGGCCAA 3480
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 3823 TGGAGAACACAGCCGGGTGTATCAGATGAAGAAATTAATTCAGAGCCAGGAGGCCAA 3882
OY 3481 CATACACACTTCATCGAGACACTCCCTGAGAAATACACACAGAGTAGAGACAAAG 3540
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 3883 CATACACTTCATCGAGACACTCCCTAATAATATACACTAAAGTAGAGACAAAG 3942
OY 3541 AACCCAGCTCTGCTGGGCGAGAAACAGCCATTCGATGCTGCGGCTTGTAGACA 3600
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 3943 AACCTAGCTCTGCTGGGCGAGAAACAAACGATTCGATGCTGCGGCTTGTAGACA 4002
OY 3601 GCGTCATATTTGCTTTGGATGAAGCTACATCAGCTCTGATACAGAAAGTGAAGGT 3660
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 4003 GCGTCATATTTGCTTTGGATGAAGCTACATCAGCTCTGATACAGAAAGTGAAGGT 4062
OY 3661 TGTCCAAAGAGCCCTGGACAAAGCCAGAGAGCCGACCTGCTGATGATGCCACCG 3720
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 4063 TGTCCAAAGAGCCCTGGACAAAGCCAGAGAGCCGACCTGCTGATGATGCCACCG 4122
OY 3721 CTTGTCCACCATCAGATGCAATGATTTAATGTTGTTTCAGAAATGGCAAGTCAAGGA 3780
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 4123 CTTGTCCACCATCAGAAATGCAACTTAATATGATGTTGTTTCAGAAATGGCAAGTCAAGGA 4182
OY 3781 GCATGGACACATCAGAAAGCTGCTGGCCGCAAGAAAGCACTATTTTCCATGAGCAGGT 3840
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 4183 GCATGGACACATCAGAAAGCTGCTGGCCGCAAGAAAGCACTATTTTCCATGAGTGAAGT 4242
OY 3841 CCAGGCTGGAGCAAGCCGTAGTGAAGTGGCCATATGAGCTGTAAATATTTTAAAT 3900
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 4243 CCAGGCTGGAGCAAGCCGTAGTGAAGTGGCCATATGAGCTGTAAATATTTTAAAT 4302
OY 3901 ATTGTGTGTTAAACATGCAATTAATCAAGTTAAAGTGAAGCTTACGAGAAACCT 3960
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 4303 ATT--TGTTTAAATNAGACATTTATTCAGATTAA--AGCAAAACCTTACAGAAAT 4355
OY 3961 ATGTAAGACCTACCTGTTTAAACATTTTGTGCGACATGAGATCTCCACCAAGTTCAG 4020
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 4356 ATGTAAGGATATCTGTTTAACTTTTCCAGTCAAGTTCAGAGCTTC--AGCAAAACCTTACAGAAAT 4403
OY 4021 AGTCTTCAAGTTTATTAATTAAGAACCA--AAGAAACATTAATCTGATGCAATTAAT 4078
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 4404 ----AGAGACTTCGTAATTAAGAACAGAGTAGAGACATCAAGTGGAGCAAT 4458
OY 4079 ACTGTGTATTAATGATATTAATAATTAAGATTAATCAAGTGAATTTGTTAATAAT 4138
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 4459 CATAGTTTAACTGATTAATAATTTTAATTAACGAATTAAGATTTTAAAGATTAATA 4518
OY 4139 TGTATTAATTTGTTTATTAATTTT-----ATTGTAATTAATGCTTTGCTGAAGAT 4190
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 4519 ATGTGATATTTGTTTATTAATTTTCCATTTGAGCTGTAAGTCTGCTGTTAAGAT 4578
OY 4191 TATAGAGTGTAAAGAAAGTACTG--ATGTTGAATTAAGTACTATATTAATACTAA 4249
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 4579 TATAGAGTGTAAAGAAAGTACTG--ATGTTGAATTAAGTACTATATTAATACTAA 4635
OY 4250 ACTTTTATATCAAAAAAAAAAAAAAAAAAAAA 4279
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 4636 ACTTTTATGTGAAAAAAAAAAAAAAAAAAAA 4665

```

RESULT 10
 ABR52041
 ID ABR52041 standard; CDNA: 4669 BP.
 XX

```

AC ABR52041:
XX
XX 13-AUG-2002 (first entry)
DT
XX
XX cDNA encoding human P-glycoprotein (Pgp) MDR1 protein.
DE
XX Human: multidrug resistance; MDR1: P-glycoprotein; Pgp: Pgp inhibitor;
KM P-glycoprotein binding; chemotherapy; cancer; cytostatic; gene; ss:
KM MRP; multidrug resistance protein.
XX
XX Homo sapiens.
OS
XX
XX Key Location/Qualifiers
XX 5'UTR 1..424
XX FT /tag= a
XX CDS 425..467
XX FT /tag= b
XX FT /product= "Human P-glycoprotein (Pgp) MDR1 (multidrug
XX FT 465..469 resistance) protein"
XX FT /tag= c
XX
XX US6365357-B1.
XX
XX 02-AER-2002.
XX
XX 21-MAY-1999; 99US-0316167.
XX
XX 15-NOV-1996; 96US-0752447.
XX
XX (ONOT-) ONOTECH INC.
XX
XX Mechtner E, Friehauf J;
XX MPI; 2002-433425/46.
XX P-PDB: AAO97493.
XX
XX Screening P-glycoprotein inhibitors by incubating cell expressing the
XX protein in presence or absence of compound, reacting cell with antibody
XX specific to protein and comparing cell- antibody binding -
XX
XX Example 1; Fig 1; 54pp; English.
XX
XX The present invention relates to a new method of screening a compound
XX for P-glycoprotein (Pgp) binding (Pgp inhibitors). The method involves
XX incubating a mammalian cell expressing Pgp in the presence or absence of
XX the compound, reacting the mammalian cell with an antibody specific for
XX Pgp in a biochemical conformation adopted in the presence of Pgp
XX substrate, and comparing binding of the antibody to the cell, in the
XX presence and absence of the compound. The method of the invention is
XX useful for screening a compound for Pgp binding, and is used in high
XX throughput screening assays to develop more effective chemotherapeutic
XX treatment of human cancer patients. The method provides information on
XX both Pgp expression and function simultaneously. Moreover, the use of
XX immunological reagents specific for Pgp reduces the possibility that the
XX assay results contain contributions from related species involved in
XX multidrug resistance, such as MRP (multidrug resistance protein). The
XX present nucleic acid sequence encodes the human P-glycoprotein (Pgp)
XX MDR1 (multidrug resistance) protein of the invention.
XX
XX Sequence 4669 BP; 1393 A; 894 C; 1130 G; 1252 T; 0 other;
XX
XX Query Match 77.9%; Score 3333.2; DB 24; Length 4669;
XX Best Local Similarity 87.5%; Pred. No. 0;
XX Matches 3753; Conservative 0; Mismatches 493; Indels 44; Gaps 8;
OY 1 GGAGCGCGAGTCTGGATGATCTCTGAAGAGCCCTAAGGGAGTGAAGAGAACTT 60
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 409 GGAAGCCGAGGTCTGGATGATCTTGAAGGAGCCCAATGAGAGCAAGAAAGAA 468
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
OY 61 CTGGAATATGGCAAAAAAAGTAAAAAATGAGAGAAAGAAAGCAACTGTCAG 120
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 469 CTTTTTAACTGAACATTAAGTAAAAAGATAGAGAAAGAAAGCAACTGTCAG 528
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||

```

QY 121 CACGTTTGCAATGTTTCGATTCATTAATTTGGCTTGATGAGTTGTATATGTTGGTGGGAC 180
DB 529 TGTATTTTCAATGTTTTCGATTCATTAATTTGGCTTGATGAGTTGTATATGTTGGTGGGAC 588
QY 181 AATGGCTCCATCCATCCATGAGTGGTCCCTCTCATGATGCTGTTTGGAAACAT 240
DB 589 TTTGGCTCCATCCATCCATGAGTGGTCCCTCTCATGATGCTGTTTGGAAACAT 648
QY 241 GACAGATAGCTTTCGAATGTCAGAAATTTTCAGAAACAAACTTTTCCAGTTATATTA 300
DB 649 GACAGATAGCTTTCGAATGTCAGAAATTTTCAGAAACAAACTTTTCCAGTTATATTA 702
QY 301 TGAAGATTTTCGAATGTCAGAAACATTTTCATCCATCCATGAGGGAATTCACAC 360
DB 703 TATATGAAAGTATTCATGATGATACAGGGTTCCTCATGATGTCAGAGAACATGACAC 762
QY 361 GTATGCTTATTTATACAGTGGGATGCGTGGCTGGTGGCTGTTACATCCAGT 420
DB 763 GTATGCTTATTTATACAGTGGGATGCGTGGCTGGTGGCTGTTACATCCAGT 822
QY 421 TTCAATTCGTGCTGTCAGACAGAAAGACAGATACCTCAAAATTAAGAAACATTTTCA 480
DB 823 TTCAATTTGGTGGCTGTCAGACAGAAAGACAGATACCTCAAAATTAAGAAACATTTTCA 882
QY 481 TGCATATTCGTCAGACAGAAAGTGGCTGGTGGTGGTGGTGGTGGTGGTGGTGGTGG 540
DB 883 TGCATATTCGTCAGACAGAAAGTGGCTGGTGGTGGTGGTGGTGGTGGTGGTGGTGG 942
QY 541 CCGGCTACAGAGATGTCCTCCAAATTCAGAAAGATGGGAGGAGGAGGAGGAGGAGT 600
DB 943 CCGGCTACAGAGATGTCCTCCAAATTCAGAAAGATGGGAGGAGGAGGAGGAGGAGT 1002
QY 601 CTTTCACTCAATAGCAATTTTTCACCGGTTTATAGTGGGGTTTACACGTGGTGA 660
DB 1003 CTTTCACTCAATAGCAATTTTTCACCGGTTTATAGTGGGGTTTACACGTGGTGA 1062
QY 661 GCTAACCTTGATTTTGGCCATCAGCCCTGTTTGGATTTGGATTTGGATTTGGATTTGG 720
DB 1063 GCTAACCTTGATTTTGGCCATCAGCCCTGTTTGGATTTGGATTTGGATTTGGATTTGG 1122
QY 721 AAAGATTCATCTTCATTCATTCATTCATTCATTCATTCATTCATTCATTCATTCATTC 780
DB 1123 AAAGATTCATCTTCATTCATTCATTCATTCATTCATTCATTCATTCATTCATTC 1182
QY 781 AGCTGAAGAGCTTTAGCAGCAATCAGAACTGTGATTTGCTTTGGAGCAAAAGAA 840
DB 1183 AGCTGAAGAGCTTTAGCAGCAATCAGAACTGTGATTTGCTTTGGAGCAAAAGAA 1242
QY 841 ACTTGAAGAGTACAAACAAATTTAGAGAGAGCTTAAGAGATTTGGGATTAAGAAAGCTAT 900
DB 1243 ACTTGAAGAGTACAAACAAATTTAGAGAGAGCTTAAGAGATTTGGGATTAAGAAAGCTAT 1302
QY 901 CAGGGCCCAACATTTCTATGTTGGTGGCTTTCTTATTTGATTCATTCATTCATTCATTC 960
DB 1303 TACAGCCCAATTTCTATGTTGGTGGCTTTCTTATTTGATTCATTCATTCATTCATTC 1362
QY 961 TTTTCTGATGGAGCTCTCTGCTCTCTCCAGTGAATATTTCTATTTGGAGAGTACTAC 1020
DB 1363 CTTTCTGATGGAGCACTCTGCTCTCTCCAGGGAATTTCTATTTGGAGAGTACTAC 1422
QY 1021 TGTCTCTTTTCTGATTTAATTTGGGCTTTAGTATTTGGAGAGGATCCCAAGATTTGA 1080
DB 1423 TGTATTTCTTTCTGATTTAATTTGGGCTTTAGTATTTGGAGAGGATCCCAAGATTTGA 1482
QY 1081 AGCATTTTGCAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 1140
DB 1483 AGCATTTTGCAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 1542
QY 1141 AAGCATTTGACATTTTGCAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 1200
DB 1543 AAGCATTTGACATTTTGCAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 1602

QY 1201 ATTCAAAATGTTCACTTCAGTTACCTTTCGAAAAGAGTTAAGATCTAAAGGCTCT 1260
DB 1603 ATTCAGAAATGTTCACTTCAGTTACCTTCAGTAATTCGAAAAGAGTTAAGATCTTAAGGCGCT 1662
QY 1261 CAACCTGAAGTTTCAGAGTGGGCAACAGTGGCGCTTTGGGAAACAGTGGCGGGA 1320
DB 1663 GAACCTGAAGTTCAGAGTGGGCAACAGTGGCGCTTTGGGAAACAGTGGCGGGA 1722
QY 1321 GAGCAGACCGTGCAGCTGATSCAGAGGCTCTATGACCCACAGATGAGTGTGTAT 1380
DB 1723 GAGCAGACCGTGCAGCTGATSCAGAGGCTCTATGACCCACAGATGAGTGTGTAT 1782
QY 1381 TGATGACAGGACATTTAGGACCATTAATGTAAAGCATCTTCGGAATTAATCTGTGTGT 1440
DB 1783 TGATGACAGGACATTTAGGACCATTAATGTAAAGTTTCTACGGGAATCATTTGTGTGT 1842
QY 1441 GAGTCAGAGGCTGT 1500
DB 1843 GAGTCAGAGGCTGT 1902
QY 1501 AAATGTCACCGATGAGATGAGATTAAGAAAGCTGTTAAGAAAGCCATGCTATGATTTAT 1560
DB 1903 AAATGTCACCGATGAGATGAGATTAAGAAAGCTGTCAGAAAGCCATGCTATGATTTAT 1962
QY 1561 CATGAACCTACCTAATTAATTTGACACTCTGTGTGAGAGAGAGGGGCCACACTGATG 1620
DB 1963 CATGAACCTACCTAATTAATTTGACACTCTGTGTGAGAGAGAGGGGCCACACTGATG 2022
QY 1621 TGGACAGAAACAGAGATGAGATGAGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 1680
DB 2023 TGGACAGAAACAGAGATGAGATGAGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 2082
QY 1681 CCTGATGAGGCAAGCTCAGCTCTGACACATGGAAGTGAAGCGTGTGTGTGTGTGTGT 1740
DB 2083 CCTGATGAGGCAAGCTCAGCTCTGACACATGGAAGTGAAGCGTGTGTGTGTGTGTGT 2142
QY 1741 GGATTAAGGCGAGAAAGGCGGACTACCATTTGTGTGTGTGTGTGTGTGTGTGTGTGT 1800
DB 2143 GGATTAAGGCGAGAAAGGCGGACTACCATTTGTGTGTGTGTGTGTGTGTGTGTGTGT 2202
QY 1801 TAATGCCGATGTATTTGT 1860
DB 2203 TAATGCCGATGTATTTGT 2262
QY 1861 TGAACCTATTAAGAAAGGAGATTTACTTCAACTGTGTGTGTGTGTGTGTGTGTGTGT 1920
DB 2263 TGAACCTATTAAGAAAGGAGATTTACTTCAACTGTGTGTGTGTGTGTGTGTGTGTGT 2322
QY 1921 TGAATTTGAGTTAGAAATGCCACTGTGTAATCCAAAGTGAAGTATGCTTGAAT 1980
DB 2323 TGAATTTGAGTTAGAAATGCCACTGTGTAATCCAAAGTGAAGTATGCTTGAAT 2382
QY 1981 GTCCTCAAAAGATTCAGGCTCAGTTTAAATTAAGAGATCAACGAGAGATATCA 2040
DB 2383 GTCCTCAAAAGATTCAGGCTCAGTTTAAATTAAGAGATCAACGAGAGATATCA 2442
QY 2041 TGCACCAAGGCGCAGAGAGAAAGCTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 2100
DB 2443 TGCACCAAGGCGCAGAGAGAAAGCTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 2502
QY 2101 TCCAGTTCTCTGTGAGAGATTTCAAGCTCACTCAATGAATGAGCTTATTTGTGT 2160
DB 2503 TCCAGTTCTCTGTGAGAGATTTCAAGCTCACTCAATGAATGAGCTTATTTGTGT 2562
QY 2161 TGTATATTTTGTGTATTAATTAAGAGGCGCTGCACACAGATTTTCAATTAATTTTTC 2220
DB 2563 TGTATATTTTGTGTATTAATTAAGAGGCGCTGCACACAGATTTTCAATTAATTTTTC 2622
QY 2221 AAGATTAATGAGATCTTTACCCGAGATGAGGATCTCGAAGCAAAAGCAGAAATATGA 2280
DB 2623 AAGATTAATGAGGCTTTTACAGAGATGATGATCTCGAAGCAAAAGCAGAAATATGA 2682
QY 2281 CATGTTTCTGTATTTGTTCATGCTTGAATTAATTTCTTTATTTACATTTTCTCTCA 2340

Db 2683 CTTGTTTACATATGTTTCTAGACCTTGTAATATTTCTTTATATACATTTTCCCTCA 2742
Qy 2241 GGGCTACATTTGGCAAGCTGGGAGATCTCTACTAGCGGCTTGATATGTTT 2400
Db 2743 GGGTTTCAATTTGGCAAGCTGGGAGATCTCTACTAGCGGCTTGATATGTTT 2802
Qy 2401 CAGATCCATGCTGAGACGATGTCAGTGGTTTGAATGACCTTAAACACACCTGAGC 2460
Db 2803 CCGATCCATGCTGAGACGATGTCAGTGGTTTGAATGACCTTAAACACACCTGAGC 2862
Qy 2461 ATTGACACCGGCTTGCAATGATGAGGCTCAAGTTAAAGGGGCTATGTTTCAAGCT 2520
Db 2863 ATTGACACCGGCTTGCAATGATGAGGCTCAAGTTAAAGGGGCTATGTTTCAAGCT 2922
Qy 2521 TGTGTCATTTACCGAATATAGCAAACTTTGGACAGCATTATTAATCTTAATCTA 2580
Db 2923 TGTGTCATTTACCGAATATAGCAAACTTTGGACAGCATTATTAATCTTAATCTA 2982
Qy 2581 TGTGTCATTTACCGAATATAGCAAACTTTGGACAGCATTATTAATCTTAATCTA 2640
Db 2983 TGTGTCATTTACCGAATATAGCAAACTTTGGACAGCATTATTAATCTTAATCTA 3042
Qy 2641 TGTGTCATTTACCGAATATAGCAAACTTTGGACAGCATTATTAATCTTAATCTA 2700
Db 3043 TGTGTCATTTACCGAATATAGCAAACTTTGGACAGCATTATTAATCTTAATCTA 3102
Qy 2701 TGGGAAATGCTACAGAAAGCCATGCAAACTTCCGAATGTTTCTTTGACTCGGGA 2760
Db 3103 TGGGAAATGCTACAGAAAGCCATGCAAACTTCCGAATGTTTCTTTGACTCGGGA 3162
Qy 2761 GCGAAGTTTGAATATATATATGACACAGATTTGCAATATACAGAACTTTGAG 2820
Db 3163 GCGAAGTTTGAATATATATATGACACAGATTTGCAATATACAGAACTTTGAG 3222
Qy 2821 GAAAGACACATCTTGGGCTCTCATTTTCTATCACCGACATGATATTTTCTTA 2880
Db 3223 GAAAGACACATCTTGGGCTCTCATTTTCTATCACCGACATGATATTTTCTTA 3282
Qy 2881 TGTGTCATTTACCGAATATAGCAAACTTTGGACAGCATTATTAATCTTAATCTA 2940
Db 3283 TGTGTCATTTACCGAATATAGCAAACTTTGGACAGCATTATTAATCTTAATCTA 3342
Qy 2941 TGTGTCATTTACCGAATATAGCAAACTTTGGACAGCATTATTAATCTTAATCTA 3000
Db 3343 TGTGTCATTTACCGAATATAGCAAACTTTGGACAGCATTATTAATCTTAATCTA 3402
Qy 3001 ATTGTCCTGACATGACCAAGCAAGTATAGCAGCCACATCATCATGATCATTTGA 3060
Db 3403 ATTGTCCTGACATGACCAAGCAAGTATAGCAGCCACATCATCATGATCATTTGA 3462
Qy 3061 AAAAAAGCTCTGATGACAGCTACAGCCCTACAGGCTCAGACCAATACGTTGGAAG 3120
Db 3463 AAAAAAGCTCTGATGACAGCTACAGCCCTACAGGCTCAGACCAATACGTTGGAAG 3522
Qy 3121 AATGTCATTTAATGAGAGTGTCTTCAATCCACTGACACAGACATCCCGTGTCT 3180
Db 3523 AATGTCATTTAATGAGAGTGTCTTCAATCCACTGACACAGACATCCCGTGTCT 3582
Qy 3181 CCAAGGCTGAGGCTCGAGGTGAAGAGGCTGAGGCTCGTGAATGAGTGAAG 3240
Db 3583 TCAAGGCTGAGGCTCGAGGTGAAGAGGCTGAGGCTCGTGAATGAGTGAAG 3642
Qy 3241 CTGTGGAGAGACAGCTTGTCAAGCTCTTGAAGGCTTCTATGACCCCTTGGGTG 3300
Db 3643 CTGTGGAGAGACAGCTTGTCAAGCTCTTGAAGGCTTCTATGACCCCTTGGGTG 3702
Qy 3301 AGTCTAATTTGATGCAAGAGATTAAGCACCCTGATGTCAGTGGTCCGACACCT 3360
Db 3703 AGTCTAATTTGATGCAAGAGATTAAGCACCCTGATGTCAGTGGTCCGACACCT 3762
Qy 3361 GGGCATGCTGCTGAGGAGCCATCTGTTGACTGACGATTCGAGAGAACTTGCCTA 3420
Db 3420 GGGCATGCTGCTGAGGAGCCATCTGTTGACTGACGATTCGAGAGAACTTGCCTA 3480

Db 3763 GGGCATGCTGCTGAGGAGCCATCTGTTGACTGACGATTCGATGAACTTGCCTA 3822
Qy 3421 TGGAGACACAGGCTGGGTGTATACATGAGAGATTTATGACGACGCAAGAGGCAA 3480
Db 3823 TGGAGACACAGGCTGGGTGTATACATGAGAGATTTATGACGACGCAAGAGGCAA 3882
Qy 3481 CATACACACTTTCATGAGACACTCCCTGAGAAATATACAAACAGAGATGAGAGCAAG 3540
Db 3883 CATACACACTTTCATGAGACACTCCCTGAGAAATATATATATATGACCTAATATGAGAGCAAG 3942
Qy 3541 AACCCAGCTCTGTTGGGCAAGAAACAGGCTGATCCATGACTGCGCTTGTGAGCA 3600
Db 3943 AACCCAGCTCTGTTGGGCAAGAAACAGGCTGATCCATGACTGCGCTTGTGAGCA 4002
Qy 3601 GCTGATATTTTGTGTTGATGATGAGTACATGCTGCTGATATACAGAAAGTAAAGT 3660
Db 4003 GCTGATATTTTGTGTTGATGATGAGTACATGCTGCTGATATACAGAAAGTAAAGT 4062
Qy 3661 TGTCCAGAGAGGCTGGGCAAGAAAGGCAAGAGGCGCAGCTGATGATGATGATGATG 3720
Db 4063 TGTCCAGAGAGGCTGGGCAAGAAAGGCAAGAGGCGCAGCTGATGATGATGATGATG 4122
Qy 3721 CTTGTCACATCCAGATGATGATGATGATGATGATGATGATGATGATGATGATG 3780
Db 4123 CTTGTCACATCCAGATGATGATGATGATGATGATGATGATGATGATGATGATG 4182
Qy 3781 GCATGAGCAACATACAGATGCTGCTGCGCCAGAAAGGCAATATTTTCCATGATGATG 3840
Db 4183 GCATGAGCAACATACAGATGCTGCTGCGCCAGAAAGGCAATATTTTCCATGATGATG 4242
Qy 3841 CCAAGCTGAGAGCAAGGCTGATGATGATGATGATGATGATGATGATGATGATGATG 3900
Db 4243 CCAAGCTGAGAGCAAGGCTGATGATGATGATGATGATGATGATGATGATGATGATG 4302
Qy 3901 ATTGTCCTGACATGACCAAGCAAGTATAGCAGCCACATCATCATGATCATTTGA 4360
Db 4303 ATTGTCCTGACATGACCAAGCAAGTATAGCAGCCACATCATCATGATCATTTGA 4355
Qy 3961 ATTGTCCTGACATGACCAAGCAAGTATAGCAGCCACATCATCATGATCATTTGA 4402
Db 4355 ATTGTCCTGACATGACCAAGCAAGTATAGCAGCCACATCATCATGATCATTTGA 4403
Qy 4021 AGTCTGATTTTAAATTAATTAAGACCA--AAGAAACATTAATGATGATGATGAT 4078
Db 4404 AGTCTGATTTTAAATTAATTAAGACCAAGTATGATGATGATGATGATGATGATG 4458
Qy 4079 ACTGCTGATTAATGATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 4138
Db 4459 ACTGCTGATTAATGATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 4518
Qy 4139 TGTATATTTTGTGTTTAAATTTT-----ATTGTAATGATGATGATGATGATGAT 4190
Db 4519 TGTATATTTTGTGTTTAAATTTTCCATTGAGTACTGATGATGATGATGATGATGAT 4578
Qy 4191 TATAGAACTGTAAGAAAGTACTG--AATGTTGAATTAAGGCTGATTAATTAATTAAT 4249
Db 4579 TATAGAACTGTAAGAAAGTACTG--AATGTTGAATTAAGGCTGATTAATTAATTAAT 4635
Qy 4250 ACTTTTAAATCAAAAAAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 4279
Db 4636 ACTTTTAAATCAAAAAAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 4665

RESULT 11
AAZ94738 standard; cDNA; 4646 BP.
AAZ94738:
01-AUG-2000 (first entry)
Human ATP binding cassette ABC1 (MDR1) cDNA.
XX

ABCB1, ATP binding cassette; human; cholesterol, lipid disorder;
atherosclerosis; lipid disorder; dyslipidemia; psoriasis;
lupus erythematosus; diagnosis; gene therapy; MRL1;
multidrug resistance; chromosome 7q21; ss.
XX Homo sapiens.
XX WO200018912-A2.
XX 06-APR-2000.
XX 21-SEP-1999; 99WO-EP06991.
XX 25-SEP-1998; 98US-0101706.
XX (FARB) BAYER AG.
XX Schmitz G, Klucken J;
XX WPI; 2000-293151/25.
XX Adenosine triphosphate binding proteins useful for identifying agents
XX for treating atherosclerosis and other inflammatory disorders
XX
XX Claim 9; Page 110-112; 154pp; English.
XX
XX The present sequence is that of human ATP binding cassette
XX subfamily B protein ABCB1 cDNA. The cDNA was identified using a
XX differential display method in which monocytes from peripheral
XX blood were subjected to macrophage differentiation and subsequent
XX loading with acetylated low density lipoproteins and subsequent
XX degrading with high density lipoprotein (HDL3) to identify
XX cholesterol sensitive genes. The gene maps to chromosome 7q21
XX and is also termed MDR1 (multidrug resistance). The invention
XX provides cholesterol-sensitive ABC genes (see AA294734-63). These
XX genes, and polypeptides encoded by them, can be used for diagnostic
XX and therapeutic applications, and for biochemical or cell-based
XX assays to screen for pharmacologically active modulator compounds
XX useful for the treatment of lipid disorders, atherosclerosis or
XX other inflammatory diseases such as psoriasis and lupus
XX erythematosus.
XX
SQ Sequence 4646 BP; 1371 A; 892 C; 1129 G; 1254 T; 0 other;
Query Match 77.7%; Score 3324.4; DB 21; Length 4646;
Best Local Similarity 87.6%; Pred. No. 0; Mismatches 486; Indels 44; Gaps 8;
Matches 3740; Conservative 0;

||||| 822
Db 763 GATGGCTATTTATACAGTGGATTTGCTGGGCTGCTGCTTACATTCAGCT
421 TTGATTTCTGTCCTGGCAGCAGAGACAGATCTCAAAATTTGAAACAAATTTTCA 480
823 TTTATTTTGGTGGCTGGCAGCTGGAACAAATACCAAAATTTGAAACAGTTTTTCA 882
481 TGTATCAGTCAGCAGAGAAATTTGGCTTGGCTGACATGCTGGGAGCTTAAAC 540
883 TGTATATGCGCAGAGAGATAGCTGTTTATGATGTGACATGCTGGGAGCTTAAAC 942
541 CCGGCTCAGCAGAGATGCTCCAAAATCAAGAGAAATTTGGCCAAATTTGAATGT 600
943 CCGACTTACAGATGATGCTCTAAGATTAATGAAATTTGCTGACAAATTTGAATGT 1002
601 CTTTCACTCAATAGCAATTTTCCACGGTTTTATAGTGGGTTTACACGCTGTGAA 660
1003 CTTTCACTCAATAGCAATTTTCCACGGTTTTATAGTGGGTTTACACGCTGTGAA 1062
661 GCTAACCTTGTGATTTTGGCCATCAGCCCTGTTGACCTTTACGCCCATCTGGGC 720
1063 GCTAACCTTGTGATTTTGGCCATCAGCCCTGTTGACCTTTACGCCCATCTGGGC 1122
721 AAGATTAATCTTCAATTTTACGTAAGAACTCTGGCTTATGAGCAATTTGAGCA 780
1123 AAGATTAATCTTCAATTTTACGTAAGAACTCTGGCTTATGAGCAATTTGAGCA 840
781 AGCTGAAGATCTTACGTAAGAACTCTGGCTTATGAGCAATTTGAGCAATTTGAGCA 1242
1183 AGCTGAAGATCTTACGTAAGAACTCTGGCTTATGAGCAATTTGAGCAATTTGAGCA 900
841 ACTTGAAGATCTTACGTAAGAACTCTGGCTTATGAGCAATTTGAGCAATTTGAGCA 960
1243 ACTTGAAGATCTTACGTAAGAACTCTGGCTTATGAGCAATTTGAGCAATTTGAGCA 1302
901 CACGGCCACATTTTATGAGTGGCTTCTTATGAGCAATTTGAGCAATTTGAGCA 1362
1303 TACAGCAATTTTATGAGTGGCTTCTTATGAGCAATTTGAGCAATTTGAGCA 1422
961 TTTCTGATGAGCTCTTGGCTTCTTATGAGCAATTTGAGCAATTTGAGCA 1480
1363 CTTCTGATGAGCTCTTGGCTTCTTATGAGCAATTTGAGCAATTTGAGCA 1540
1021 TGTCTTCTTCTGATTTATGAGTGGCTTCTTATGAGCAATTTGAGCAATTTGAGCA 1580
1423 TGTATTTCTTCTGATTTATGAGTGGCTTCTTATGAGCAATTTGAGCAATTTGAGCA 1640
1081 AGCATTTGCAACGAG 1700
1483 AGCATTTGCAACGAG 1760
1141 AAGCATTTGCAACGAG 1820
1543 AAGCATTTGCAACGAG 1880
1201 ATTCAAAATTTTCACTTCACTTCACTTCACTTCACTTCACTTCACTTCACTTCACT 1940
1603 ATTCAAAATTTTCACTTCACTTCACTTCACTTCACTTCACTTCACTTCACTTCACT 2000
1261 CAACTGAAGGTGAGAGTGGGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1722
1663 GAACTGAAGGTGAGAGTGGGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1782
1321 GAGCAG 1842
1723 GAGCAG 1902
1381 TGTATGAG 1962
1783 TGTATGAG 2022
1441 GAGTCAAG 2082

QY	2561	TGTTGGCAATTAACACTTTTACTCTTAGCAATGTGACCATCATTTGCAATTACAGAGT	264
Db	2983	TGTTGGCACTAACACACTGTTACTTAGCAATTGTACCATCATATTGACATTCAGAGAT	304
QY	2641	TGTTGAATGAAATGTGTGTCGACACAGCACTGAAAGATTAAGAAAGAGCTGAAGAGAC	2700
Db	3043	TGTTGAATGAAATGTGTGTCGACACAGCACTGAAAGATTAAGAAAGAGCTGAAGAGTGC	3100
QY	2701	TGGAGAAATTGCTAACAGAAACCATCGAAACTCCGAACTGGTGTCTTTACACGCGGA	2760
Db	3103	TGGAGAAATGCTAACTGAACCAATAGAAACTTCGAAACCGTTGTTCTTTACTACAGAA	3160
QY	2761	GCAGAAATTTGAATACATGATATGATCCAGAGTTTGCAAGTACCATACAGAACTCTTGAG	2820
Db	3163	GCAGAAATTTGAACATATGATATGATCTCAGAGTTTGAGAGTACCATACAGAACTCTTGAG	3222
QY	2821	GAAGACACACATCTTCGGGGGTCTCATTTTCTATCACCCAGGCAATGATGTATTTTCTCTA	2880
Db	3223	GAAGACACACATCTTTGGAAATTAACATTTTCTTCTTCCACCGGAATGATGTATTTTCTCTA	3282
QY	2881	TGCTGCTCTTTCCGGTTTGGTGCCACTGGTGGCCAAATAGTTCAATGAACCTTTCAGGA	2940
Db	3283	TGCTGATGTCTTCGGTTTGGAGCCTACTTGGTGGCAATTAACCTCAATGACCTTTTGAGGA	3342
QY	2941	TGTTCTTTTGGATTCCTACGCTATTGTCTTTGGTGCCATGGCAGTGGGGCAGGCTATTC	3000
Db	3343	TGTTCTGTGATGATTTTTCAGCTGTGTGTTGGTGCCATGGCCGTGGGGCAAGTCAATTC	3402
QY	3001	ATTGTGCTCTGCACTATGCCAAAGCCAAAGTATACAGAGCCGACGCTCATATGATCATTTGA	3060
Db	3403	ATTGTGCTCTGCACTATGCCAAAGCCAAAGTATACAGAGCCGACGCTCATATGATCATTTGA	3462
QY	3061	AAAAAGCCCTCGATTTGACCTACAGCCCTCACAGCCTCAAGCCAAATACGTTGGAAGG	3120
Db	3463	AAAAAGCCCTTTGATTTGACCTACAGCAGGAGGCGCTATATCCGACACATTTGGAAGG	3522
QY	3121	AAATGTGACATTTATGAGGTCGTGTCAACTATCCCATCGACAGACATATCCCGCTGT	3180
Db	3523	AAATGTGACATTTGATGAGGTTGATTAATCACTATCCACCCGACCGACATCCCACTGT	3582
QY	3181	CCAGGGGCTGAGCCTCGAGGTGAAGAGGCGCCAGACGCTGGCCTGTAGTAGCAGTTGG	3240
Db	3583	TCAGGAGCTAGCCTGAGGAGTGAAGAGGCGCCAGACGCTGGCTGTGTGGGACGAGTTGG	3642
QY	3241	CTGTGGGAAGACACAGTTTTCAGTCTCCAGAGCCCTCTATGACCCCTTGGGCTGTTCC	3300
Db	3643	CTGTGGGAAGACACAGTGTGTCAGGCTCCGAGCGGTTCTACGACCCCTTGGCAGGGGAA	3702
QY	3301	AGTCTAATTGATGGCAAGAGATTAAGACCCGAAATGTCCAGTGCTCCAGACACACT	3360
Db	3703	AGTCTCTCTTGATGGCAAGAAATTAAGGACGATGTCTAGTGCTCCAGGACACACT	3762
QY	3361	GGGCAATCTGTCTCAGAGAGCCATCTCTGTTTGAATGACACATTTGCCAGAACATTTGCTTA	3420
Db	3763	GGGCAATCTGTCTCAGAGAGCCATCTCTGTTTGAATGACACATTTGCCAGAACATTTGCTTA	3822
QY	3421	TGGAGACAAACAGCCGGGTGTATCATCATTAAGATTTTGGACGAGAGCCAGAGGCCAA	3480
Db	3823	TGGAGACAAACAGCCGGGTGTCTACAGAGAGAGTCTGTAGGGGCAAGAAAGAGGCCAA	3882
QY	3481	CATACACCACTTCATCAGACACTCCCTGAGAAATATCAACAACAGAGATGAGAGCAAGG	3540
Db	3883	CATACATGCTTCATCAGACTCCTCCCTTAATTAATATAGCACTAAATGAGAGCAAGG	3942
QY	3541	AAACCAAGTCTCTGTGGCCAGAAACACAGGATTTGCCATGAGCTCGGCTCTTGTAGACA	3600
Db	3943	AACTCAGCTCTCTGTGGCCAGAAACACAGATTTGCCATGAGCTCGGCTCTTGTAGACA	4002
QY	3601	GGCTATATTTTGGCTTTTGGATGTAAGTACATCAGCTCTGGATACAGAAATGTAAGGT	3660
Db	4003	GGCTATATTTTGGCTTTTGGATGTAAGGACAGCTCAGCTCTGATACAGAAATGTAAGGT	4062

RESULT	12
AD38994	
ID	AD38994 standard; DNA; 4646 BP.
XX	
AC	AD38994;
XX	
DT	23-SEP-2002 (first entry)
XX	
DE	Human mdrl gene.
XX	
KW	Human; haematologic malignancy; multidrug resistance; MDR; SUMO-1;
KW	hypoxia inducible factor-1; small ubiquitin-like-modifier; HIF-1;
KW	lymphoid disorder; chronic lymphoproliferative disorder; lymphoma;
KW	myeloid disorder; lymphocytic leukaemia; thrombocythaemia; myeloma;
KW	angiogenic myeloid metaplasia; myeloid leukaemia; gene therapy;
KW	polycythaemia vera; hypoxia responsive element; HRE; gene; ds.
XX	
OS	Homo sapiens.
XX	
XX	
FT	Key
FT	CDS
FT	425..4267
FT	/tag= a
FT	/product= "Human MDR protein"
XX	
PN	WO200234291-A2.
XX	
PD	02-MAY-2002.
XX	

XX The invention relates to a method of treating a subject having or at
CC risk of developing a haematologic malignancy or multidrug resistance
CC (MDR). The method involves administering hypoxia inducible factor-1
CC (HIF-1) binding molecules or small ubiquitin-like-modifier (SUMO)-1
CC binding molecules or HIF-1-SUMO-1 complex modulators. mdrl hypoxia
CC responsive element (HRE) binding molecules or antisense nucleic
CC acid molecules and SUMO-1 binding molecules or antisense molecules
CC are useful for treating a subject having or at risk of developing
CC haematologic malignancy or MDR (e.g., a lymphoid disorder or a myeloid
CC disorder). The lymphoid disorders include Lymphocytic leukaemia or
CC chronic lymphoproliferative disorders e.g., lymphoma, myeloma or acute
CC lymphoid leukaemia. The myeloid disorders include chronic or acute
CC myeloid leukaemia, e.g., angioleukemia vera. The invention is used in gene
CC therapy. The present sequence is human mdrl gene.
XX

Sequence 4646 BP: 1371 A; 892 C; 1129 G; 1254 T; 0 other;

[illegible]

Db 883 TCGTAAATGCGACAGAGATAGGCTGGTTGATGTGACATGTTGGGAGCTTAACAC 942
QY 541 CCGGCTCAGACAGATGCTCTCCAAATCATGAAGAAATTTGGCACAAAATTTGGAATGT 600
Db 943 CCGATTTACAGATGATGCTCTTAAGATTAAGAAATTTAGTGTGCAAAAATTTGGAATGT 1002
QY 601 CTTCACATCAATAGCAACATTTTACCGGGTTTATAGTGGGGTTTACAGCTGGTTGAA 660
Db 1003 CTTCAGTCAATAGCAACATTTTACAGCTGGTTTATAGTGGGGTTTACAGCTGGTTGAA 1062
QY 661 GCTAACCCCTGGATTTTGGCCATCAGCCCTGTTTGGACCTTTCAGCCGCCATCGGGC 720
Db 1063 GCTAACCCCTGGATTTTGGCCATCAGCCCTGTTTGGACCTTTCAGCTGGCTGTCTGGCC 1122
QY 721 AAAGATCTATCTTCTTACTGTATGAAGAACTCTGGCCATGACAAAAGCTGGACAGT 780
Db 1123 AAAGATCTATCTTCTTACTGTATGAAGAACTCTTGAAGTATGCAAAAAGCTGGAGCAGT 1182
QY 781 AGCTGAAGAGTCTTACAGCAATCAGAACTGATTTGCCCTTGGAGCAAAAGAAAGA 840
Db 1183 AGCTGAAGAGTCTTGGCACCAATTAGAACTGATTTGCATTTGGAGCAAAAGAAAGA 1242
QY 841 ACTTGAAGGTACACAAAATTTAGAAAGCTAAAGAAATTTGGATTAAGAAAGCTAT 900
Db 1243 ACTTGAAGGTACACAAAATTTAGAAAGCTAAAGAAATTTGGATTAAGAAAGCTAT 1302
QY 901 CAGGCCACATTTCTATTTGGTGGCCCTTTCTTATTTCTATGATCATATGCTGGC 960
Db 1303 TACAGCAATTTCTATTTAGTGTGCTGCTTCTGCTGATCATATGATTTTGTCTGGC 1362
QY 961 TTTCTGTATGGACCTCTTGTGCTCTCCAGTAAATTTCTATTTGACAAGTACTAC 1020
Db 1363 CTTCGTGATGGAGCACCTTGTCTCTCAGGGGAAATTTCTATTTGACAAGTACTAC 1422
QY 1021 TGTCTTCTTCTTGTATTAATTTGGGGCTTTTATTTGATGAGAGCAATCCCAACATTTGA 1080
Db 1423 TGTATTTCTTCTGTATTAATTTGGGGCTTTTATTTGATGAGAGCAATCCCAACATTTGA 1482
QY 1081 ASCATTTGCAAAAGCAAGAGAGAGAGCTTATGAATCTTCAAGTAAATTTGACAATTAAC 1140
Db 1483 ASCATTTGCAAAATGCAAGAGAGAGAGCTTATGAATCTTCAAGTAAATTTGACAATTAAC 1542
QY 1141 AAGCATTAACAGTATTTGAAAGTGTGACATTAACAGATTAATTTAAGGAAATTTTGA 1200
Db 1543 AAGTATTAACAGTATTTGAAAGTGTGAGAGTGGCAAAACAGATTAATTTAAGGAAATTTTGA 1602
QY 1201 ATTCAAAAATGTCATCTCACTTACCTTCTCGAAAAGAAATTTAAGATCTTAAAGGCTC 1260
Db 1603 ATTCAAAAATGTCATCTCACTTACCTTCTCGAAAAGAAATTTAAGATCTTAAAGGCTC 1662
QY 1261 CAACCTGAAGGTTGAGAGTGGGCGAGACAGTGGGCGTGGTTGGGAAAGAGTGGCTGCGGGA 1320
Db 1663 GAACCTGAAGGTTGAGAGTGGGCGAGACAGTGGGCGTGGTTGGGAAAGAGTGGCTGCGGGA 1722
QY 1321 GAGCAGACCGTGTGAGTGTATGACAGAGCTTATGAACCCACAGATGAGCTGTGTAT 1380
Db 1723 GAGCAGACCGTGTGAGTGTATGACAGAGCTTATGAACCCACAGATGAGCTGTGTAT 1782
QY 1381 TGATGACAGAGATTAAGCACTAATATGAAGGCACTCTTGGGAAATTTACTGGTGTGT 1440
Db 1783 TGATGACAGAGATTAAGCACTAATATGAAGGCACTCTTGGGAAATTTACTGGTGTGT 1842
QY 1441 GAGCAGAGGCGCTGTGTGTGTCACACAGATAGCTGAAAACATTTGCGCTATGCGCGTGA 1500
Db 1843 GAGCAGAGGCGCTGTGTGTGTCACACAGATAGCTGAAAACATTTGCGCTATGCGCGTGA 1902
QY 1501 AAATGTACCATGATGAGATTAAGAAAGCTTTAAGAAAGCAATGGCTTATGATTTAT 1560
Db 1903 AAATGTACCATGATGAGATTAAGAAAGCTTTAAGAAAGCAATGGCTTATGATTTAT 1962
QY 1561 CATGAAGCACTAATTAATTTGACATCTGTGTTGAGAGAGAGGCGCCAGCTGAGTGG 1620

Db 1963 CATGAAGCACTAATTAATTTGACACCCGTGGTGGAGAGAGAGGCGCCAGTTGATGG 2022
QY 1621 TGGACAGAAACAGAGATGCGCATTTGCTCGGGCCCTGGTTGCAACCCCAAGATTTCTG 1680
Db 2023 TGGGCAAGAGAGAGATGCGCATTTGCAAGTGGCCCTGGTTGCAACCCCAAGATTTCTG 2082
QY 1681 GCTGATAGGCAAGTGTACGCTGTGACACTGAAAGTGAAGCAAGTGGTTGAGTGGCCCT 1740
Db 2083 GCTGATAGGCAAGTGTACGCTGTGACACTGAAAGTGAAGCAAGTGGTTGAGTGGCCCT 2142
QY 1741 GGATTAAGGCCAGAAAAGGCGCACTACATTTGATAGTATGATGATGATGATGATGATG 1800
Db 2143 GGATTAAGGCCAGAAAAGGCGCACTACATTTGATAGTATGATGATGATGATGATGATG 2202
QY 1801 TTAATGCGAGTCACTGCTGCTGTTTATGATGATGATGATGATGATGATGATGATGATG 1860
Db 2203 TTAATGCGAGTCACTGCTGCTGTTTATGATGATGATGATGATGATGATGATGATGATG 2262
QY 1861 TGAACATCATGAAGAGAGAGGCAATTTACTTCAAACTTGTCACAATGCAAGCAAGAGAA 1920
Db 2263 TGAACATCATGAAGAGAGAGGCAATTTACTTCAAACTTGTCACAATGCAAGCAAGAGAA 2322
QY 1921 TGAATTTGAGTTGAAAATGCCATGCTGTGATATCCAAAAGTGAAGTGAATGCCTTGAAAT 1980
Db 2323 TGAATTTGAGTTGAAAATGCCATGCTGTGATATCCAAAAGTGAAGTGAATGCCTTGAAAT 2382
QY 1981 GTCTCCAAAAGATTTAGGGTCCAGTTTAAATAAAGAAAGATCACTGCGACAGATPACA 2040
Db 2383 GTCTCCAAAAGATTTAGGGTCCAGTTTAAATAAAGAAAGATCACTGCGACAGATPACC 2442
QY 2041 TGGACACAAAGCCCAAGACAGAAAGCTTGTACAAAAGAGACATGATGAAATGATACC 2100
Db 2443 TGGATCCAAAGCCCAAGACAGAAAGCTTGTACAAAAGAGACATGATGAAATGATACC 2502
QY 2101 TCCAGTTTCTCTGAGAGATTTGAAAGCTGAACTCAACTCAATATGCTTATTTGTGT 2160
Db 2503 TCCAGTTTCTCTGAGAGATTTGAAAGCTGAACTCAACTCAATATGCTTATTTGTGT 2562
QY 2161 TGGTATTTTGTGCTATTAATTAACAGAGGCTTCAACAGCAATTTTCATTAATTTTC 2220
Db 2563 TGGTATTTTGTGCTATTAATTAACAGAGGCTTCAACAGCAATTTTCATTAATTTTC 2622
QY 2221 AAGCATTTAAGGATCTTTACCGGAGATGAGATGCCGAAACAAAAGCAGAAATGATA 2280
Db 2623 AAGCATTTAAGGATCTTTACCGGAGATGAGATGCCGAAACAAAAGCAGAAATGATA 2682
QY 2281 CATGTTTCTGTATTTGTTTCTAGTCTTGAATTAATTTCTTATTAATTAATTTCTCTCA 2340
Db 2683 CATGTTTCTGTATTTGTTTCTAGTCTTGAATTAATTTCTTATTAATTAATTTCTCTCA 2742
QY 2341 GGGCTTCAATTTGGCAAAAGCTGGGAGATCCCTCACATAAGCGGCTTGCATATGATTTT 2400
Db 2743 GGGCTTCAATTTGGCAAAAGCTGGGAGATCCCTCACATAAGCGGCTTGCATATGATTTT 2802
QY 2401 CAGATTCATCTGAGACAGAGATGTCAGTGTGTTTGAAGCCCTTAAAAAACACACAGGAGC 2460
Db 2803 CAGATTCATCTGAGACAGAGATGTCAGTGTGTTTGAAGCCCTTAAAAAACACACAGGAGC 2862
QY 2461 ATTGAACACAGGCTTCCCAATGATGATGCGCTCAAGTTAAAGGGCTATAGTTCCAGGCT 2520
Db 2863 ATTGAACACAGGCTTCCCAATGATGATGCGCTCAAGTTAAAGGGCTATAGTTCCAGGCT 2922
QY 2521 TGTGTCAATTAACCAAAATATAGCAAAATCTTGGACAGCAATTAATTAATCTTATCTA 2580
Db 2923 TGTGTCAATTAACCAAAATATAGCAAAATCTTGGACAGCAATTAATTAATCTTATCTA 2982
QY 2581 TGTGTCAATTAACCAATTAATCTTGTAGCAATTTGACCAATTCATATGACAGAGAT 2640
Db 2983 TGTGTCAATTAACCAATTAATCTTGTAGCAATTTGACCAATTCATATGACAGAGAT 3042
QY 2641 TGTGTCAATTAATTAATTTGACATCTGTGTTGAGAGAGAGGCGCCAGCTGAGTGG 2700
Db 3043 TGTGTCAATTAATTAATTTGACATCTGTGTTGAGAGAGAGGCGCCAGCTGAGTGG 3102

Young PE, Augustus M, Carter KC, Ebner R, Endress G, Horrigan S,
Soppet DR, Weaver Z;
WPI; 2002-188264/24.

Claim 1; SEQ ID 6929; 44pp; English.

The present invention describes a method (M1) for screening for an anti-neoplastic agent. The method involves exposing cells to a chemical agent to be tested for anti-neoplastic activity, determining a change in expression of at least one gene (I) of a signature gene set, where (I) comprises a sequence (S) selected from 8447 sequences (given in AB161664 to AB170110), or is at least 95% identical to (S), where a change in expression is indicative of anti-neoplastic activity. (I) has cytostatic activity and can be used in gene therapy. M1 can be used for screening an anti-neoplastic agent, and can be used for producing a product which is the data collected with respect to the anti-neoplastic agent as a result of M1, and the data is sufficient to convey the chemical structure and/or properties of the agent. M1 can be used in the treatment of cancer such as colon, breast, stomach, lung, thyroid, oesophageal, ovarian, kidney, prostate or pancreatic cancer, adenocarcinoma, carcinoma, clear cell cancer, infiltrating ductal cancer, infiltrating lobular cancer, squamous cell carcinoma, neuroendocrine carcinoma, papillary carcinoma and Wilms' tumour.

Sequence 4646 BP; 1371 A; 892 C; 1129 G; 1254 T; 0 other;

Query Match	77.7%;	Score 3324.4;	DB 24;	Length 4646;
Best Local Similarity	87.6%;	Pred. No. 0;		
Matches 3740; Conservative	0.0;	Mismatches 406		

		missmatches	48b;	Indels	44;	Gaps
QY	1	GGAGCCGAGG	GTGGGATG	GGATCCGAAAGG	AGGCGCTTAAGGGGAGTGC	GAGAGAAATT 60
Db	409	GGAGCCGAGG	GTGGGATG	GGATCTTGAAAGGGGAGC	CGCAATGAGAGGCAAGGAAGAA	468
QY	61	CTGGAATG	GGCCAAA	AAAAATGA	AAAGAAAGAAAGAAACCACTGCAG	120

Db	469	CTTTTAAACGTGAMCAATATAAGTAAGAAAAAGATATAGAAGAAAACAAACCACACTGCTAG	532
Oy	121	CACGTTTGCAATGTTTCGGTATTCCAAATGGCGCTGATAGGTTATATATGTTGGGGAC	180
Db	529	TGATATTTCGAATGTTTCGGTATTCCAAATGGCGCTGACAAAGTTGATATATGTTGGGGAAC	588
Oy	181	AATGGCTGCCATCATCATTGAGAGCTGCACCTCCCTCATCATGATGATGCTGTTTGGGAACAT	240
Db	589	TTTGGGCTGCATCATCATTGAGAGCTGCACCTCCCTCATCATGATGATGCTGTTTGGGAACAT	648
Oy	241	GACAGATACCTTTTGGAAATGAGGAATTTAACAACAAACCTTTTCCAATTATATATTA	300
Db	649	GACAGATATCTTTGCAAATATGAGSAATTTTAGAAGATCGAATGTA-----AACATCAC	702
Oy	301	TGAAAGTATTACGAACAATATACACAACTTTTCATCAACCATCTGGAGAGAAATGACAC	360
Db	703	TATAGAAAGTATGATATCATATGATACAGGGTTCCTCATGAATCTGGAGGAACATGACAC	762
Oy	361	GATAGGCTATTATATACAGTGGGATCGGTGCTGGCGCTGCTGGTGGCTGCTTACATCCAGT	420
Db	763	GATAGGCTATTATATACAGTGGGATCGGTGCTGGCGCTGCTGGTGGCTGCTTACATCCAGT	822
Oy	421	TTTCATTTGCTGCTCGGCACAGAGAACAGATACATACTCAAAATTTGAAACAAATTTTTTA	480
Db	823	TTTCATTTTGGTGGCTCGGCACAGTGGAGAACAAATATACAAATTTGAAACAAATTTTTTA	882
Oy	481	TGCTATCATGAGSACAGAGATTGGCTGGTTTACGTCGATACGTTTGGGAGCTTATAC	540
Db	883	TGCTATCATGAGSACAGAGATTGGCTGGTTTACGTCGATACGTTTGGGAGCTTATAC	942
Oy	541	CCGGCTCACAGAGAGATGCTCCAAAATCAATGAAGAAATTGGGACAAAATTTGGAATGT	600
Db	943	CCGACTTACAGAGATGCTCCAAAATCAATGAAGAAATTGGGAGCTTATGGAATGT	1002
Oy	601	CTTTCACATCAATGCAACAATTTTACCGGTTTATAGTGGGGTTTACACGTTGTGAA	660
Db	1003	CTTTCACATCAATGCAACAATTTTACCGGTTTATAGTGGGGTTTACACGTTGTGAA	1062
Oy	661	GCTAACCCCTGTGATTTTGGCCATCAGCCGCTTGGAGCTTTCAGCCGCATATGGGC	720
Db	1063	GCTAACCCCTGTGATTTTGGCCATCAGCCGCTTGGAGCTTTCAGCCGCATATGGGC	1122
Oy	721	AAAGATATATCTCTATTTACTGATTAAGAACTCTTGGCCATATGCAAAAAGCTGAGCAGT	780
Db	1123	AAAGATATATCTCTATTTACTGATTAAGAACTCTTACGATATGCAAAAAGCTGAGCAGT	1182
Oy	781	AGCTGAGAAGCTTTAGCAGCAATCAACATGATAGCTTGGAGACAAAGAAAGA	840
Db	1183	AGCTGAGAAGCTTTAGCAGCAATTTAGACACTGATATGATTTGGAGACAAAGAAAGA	1242
Oy	841	ACGTGAAGGATACAAACAAATTTAGAAAGCTTAAGGAATTTGGATTAAGAAAGCTAT	900
Db	1243	ACGTGAAGGATACAAACAAATTTAGAAAGCTTAAGGAATTTGGATTAAGAAAGCTAT	1302
Oy	901	CACGGCAACATTTCTATGTTGGCGCGTTTCTTATGATCATGATCATATGCTGCGC	960
Db	1303	TACACCCCAATTTCTATGATGCTGCTTCTGCTGATCATGATCATATGCTGCGC	1362
Oy	961	TTTTCTGATATGGGACCTCTTGCTCTCTCCAGTGAATATTTTGGAAAGTACTCAC	1020
Db	1363	TTTTCTGATATGGGACCTCTTGCTCTCTCCAGTGAATATTTTGGAAAGTACTCAC	1422
Oy	1021	TGTTCTTTTCTGATATTAATTTGGGCTTTTATGATTTGGAGACAGCATCCCCAAGATTGA	1080
Db	1423	TGTTCTTTTCTGATATTAATTTGGGCTTTTATGATTTGGAGACAGCATCCCCAAGATTGA	1482
Oy	1081	AGCATTTTGCAACGCAAGAGAGACGCTTATGAATCTTCAAGATATTTGACATAAAC	1140
Db	1483	AGCATTTTGCAACGCAAGAGAGAGCAGCTTATGAATCTTCAAGATATTTGATATATACC	1542
Oy	1141	AAGCATTTGACAGTATTTGCAAGAGAGCAATTAACCGAGATATATTTAAGGAAATTTGA	1200
Db	1543	AAGTATTGACAGTATTTGCAAGAGTGGGCAACAAACGAGATATATTTAAGGAAATTTGA	1602

1201 ATTCAAAAATGTTCACTTCACTTACCTTCGCGAAAAAGATTAAAGATCTTAAAGGCTCT 1260
1603 ATTCAGAAATGTTCACTTCACTTACCTTCGCGAAAAAGATTAAAGATCTTAAAGGCTCT 1662
1261 CAACCTGAAGGTTTCAGAGTGGGCGAGACAGTGGCGTGGTGGGAACAGTGGCGGGGAA 1320
1663 GAACCTGAAGGTTTCAGAGTGGGCGAGACAGTGGCGTGGTGGGAACAGTGGCGGGGAA 1722
1321 GACACAGACCGTTCAGAGTGGGCGAGACAGTGGCGTGGTGGGAACAGTGGCGGGGAA 1380
1723 GACACAGACCGTTCAGAGTGGGCGAGACAGTGGCGTGGTGGGAACAGTGGCGGGGAA 1782
1381 TGATGACAGAGCATTTAGAGACCATTAATGTAAAGCATCTTTCGGGAATTTACTGGTGGT 1440
1783 TGATGACAGAGCATTTAGAGACCATTAATGTAAAGCATCTTTCGGGAATTTACTGGTGGT 1842
1441 GAGTCAGAGCGCTGTGTGTTGGCAGACAGTACCTGAAAAACATGCGTATGCGCCGGA 1500
1843 GAGTCAGAGCGCTGTGTGTTGGCAGACAGTACCTGAAAAACATGCGTATGCGCCGGA 1902
1501 AAATGTCACCATGGATGAGATGAGAAAGCTGTAAAGAGCCATGCGCTATGATTTTAT 1560
1903 AAATGTCACCATGGATGAGATGAGAAAGCTGTAAAGAGCCATGCGCTATGATTTTAT 1962
1561 CATGAACCTACCTATTAATTTGACACCTGGTGGAGAGAGAGGCGCCAGCTGAGTGG 1620
1963 CATGAACCTACCTATTAATTTGACACCTGGTGGAGAGAGAGGCGCCAGCTGAGTGG 2022
1621 TGAGAGAGAAAGAGAAATGCGCATTTGCTGGGCGCTGTGTGCGAACCCCAAGATCTTCT 1680
2023 TGAGAGAGAAAGAGAAATGCGCATTTGCTGGGCGCTGTGTGCGAACCCCAAGATCTTCT 2082
1681 GCTGATGAGCAACGTCAGCTCGTGGGACCTGAAAGAGAGAGTGGTTCAGGTGGCCCT 1740
2083 GCTGATGAGCAACGTCAGCTCGTGGGACCTGAAAGAGAGAGTGGTTCAGGTGGCCCT 2142
1741 GGATGAGCGCGAAAGAGCGCGACTACCATTTGATGATGATGATGATGATGATGATGAT 1800
2143 GGATGAGCGCGAAAGAGCGCGACTACCATTTGATGATGATGATGATGATGATGATGAT 2202
1801 TAATGCCAGTGTCAATGCTGTGTTGATGATGATGATGATGATGATGATGATGATGATGAT 1860
2203 TAATGCCAGTGTCAATGCTGTGTTGATGATGATGATGATGATGATGATGATGATGATGAT 2262
1861 TGAACTCATGAAAGAGAAAGGCGATTACTCAAACTTGTCACAAATGACAGACAGGAAA 1920
2263 TGAACTCATGAAAGAGAAAGGCGATTACTCAAACTTGTCACAAATGACAGACAGGAAA 2322
1921 TGAATTTGAGTTAGAAATGCGATGCGTGAATCCAAAAGTGAAGTATGATGCTTGGAAAT 1980
2323 TGAATTTGAGTTAGAAATGCGATGCGTGAATCCAAAAGTGAAGTATGATGCTTGGAAAT 2382
1981 GTCGCCAAAAGATTTCAGGGTCCAGTTTAAATAAAAAGAAATCAACTGCGAGAGATTAACA 2040
2383 GTCGCCAAAAGATTTCAGGGTCCAGTTTAAATAAAAAGAAATCAACTGCGAGAGATTAACA 2442
2041 TGCACCAAGAGCCAGACAGAAAGCTTGTGTACAAAAGAGAGGCTGAATGAGATGTACC 2100
2443 TGCATCACAAAGCCAGACAGAAAGCTTGTGTACAAAAGAGAGGCTGTGATGAAAGTATACC 2502
2101 TCCAGTTTCCCTTCGAGAGATCTGAACTGAACTCAACTGAATGGCTTATTTTGTGTGT 2160
2503 TCCAGTTTCCCTTCGAGAGATCTGAACTGAACTCAACTGAATGGCTTATTTTGTGTGT 2562
2161 TGGATATATTTTGTCTATTAATAAGAGAGGCTGCAACAGAGATTTTCAATTAATTTTC 2220
2563 TGGATATATTTTGTCTATTAATAAGAGAGGCTGCAACAGAGATTTTCAATTAATTTTC 2622
2221 AAGGATTTTAGGATCTTTACCGAGATGAGATCTTGAACCAAAAGCAGAGATAGTAA 2280
2623 AAGGATTTTAGGATCTTTTACAGAGATGATGATGATGATGATGATGATGATGATGAT 2682

2281 CATGTTTCTGATTTGTTTCTAGTCTTGGAAATTAATTTCTTTTATTAATTTTCTCCCA 2340
2683 CTGTTTCTAGATTTGTTTCTAGTCTTGGAAATTAATTTCTTTTATTAATTTTCTCCCA 2742
2341 GGGCTTCACATTTGGCAAAAGCTGGGAGATCTCATAAGCGGCTTGCATATGTTTTT 2400
2743 GGGCTTCACATTTGGCAAAAGCTGGGAGATCTCATAAGCGGCTTGCATATGTTTTT 2802
2401 CAGATCCATGCTGAGACAGAGATGACGCTGTTTATGATGATGATGATGATGATGATGAT 2460
2803 CAGATCCATGCTGAGACAGAGATGAGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGG 2862
2461 ATTGACACAGAGGCTGCCAATGATGCGCTCAAGTATTAAGGGCTATGATGATGATGAT 2520
2863 ATTGACACAGAGGCTGCCAATGATGCGCTCAAGTATTAAGGGCTATGATGATGATGATGAT 2922
2521 TGTGTCATTTACCGAGATTAATGCAAAATCTTGGGAGAGGATTAATTAATTAATTAATTA 2580
2923 TGTGTCATTTACCGAGATTAATGCAAAATCTTGGGAGAGGATTAATTAATTAATTAATTA 2982
2581 TGTGTCATTTACCGAGATTAATGCAAAATCTTGGGAGAGGATTAATTAATTAATTAATTA 2640
2983 TGTGTCATTTACCGAGATTAATGCAAAATCTTGGGAGAGGATTAATTAATTAATTAATTA 3042
2641 TGTGTCATTTACCGAGATTAATGCAAAATCTTGGGAGAGGATTAATTAATTAATTAATTA 2700
3043 TGTGTCATTTACCGAGATTAATGCAAAATCTTGGGAGAGGATTAATTAATTAATTAATTA 3102
2701 TGGGAAGATGCTGACAGAGCATGCAAAATCTTGGGAGAGGATTAATTAATTAATTAATTA 2760
3103 TGGGAAGATGCTGACAGAGCATGCAAAATCTTGGGAGAGGATTAATTAATTAATTAATTA 3162
2761 GGAGAAGTTTGAATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 2820
3163 GGAGAAGTTTGAATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 3222
2821 GAAAGACACATCTTGGGAGGCTCAATTTCTATGACCCAGGCAATGATGATGATGATGAT 2880
3223 GAAAGACACATCTTGGGAGGCTCAATTTCTATGACCCAGGCAATGATGATGATGATGAT 3282
2881 TGTGTCATTTACCGAGATTAATGCAAAATCTTGGGAGAGGATTAATTAATTAATTAATTA 2940
3283 TGTGTCATTTACCGAGATTAATGCAAAATCTTGGGAGAGGATTAATTAATTAATTAATTA 3342
2941 TGTGTCATTTACCGAGATTAATGCAAAATCTTGGGAGAGGATTAATTAATTAATTAATTA 3000
3343 TGTGTCATTTACCGAGATTAATGCAAAATCTTGGGAGAGGATTAATTAATTAATTAATTA 3402
3001 ATTGTCATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 3060
3403 ATTGTCATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 3462
3061 AAAAAAGCCCTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 3120
3463 AAAAAAGCCCTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 3522
3121 AAATGTCATTTAATGAGTGTGTTCAATCAATGCAATGCAATGCAATGCAATGCAATGCAAT 3180
3523 AAATGTCATTTAATGAGTGTGTTCAATCAATGCAATGCAATGCAATGCAATGCAATGCAAT 3582
3181 CCAGGGCTGAGCTGAGAGTGAAGAGAGGCGAGAGGCTGAGAGGCTGAGAGGCTGAGAGG 3240
3583 CCAGGGCTGAGCTGAGAGTGAAGAGAGGCGAGAGGCTGAGAGGCTGAGAGGCTGAGAGG 3642
3241 CTTGTCGAGAGGAGCAAGTGTGTTCAATCAATGCAATGCAATGCAATGCAATGCAATGCAAT 3300
3643 CTTGTCGAGAGGAGCAAGTGTGTTCAATCAATGCAATGCAATGCAATGCAATGCAATGCAAT 3702
3301 AGTGTCAATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 3360
3703 AGTGTCAATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 3762
3361 GGGCATGCTGTCTCAGAGGAGGCGGATGCTGTTGATGATGATGATGATGATGATGATGAT 3420

```

Db 3763 GGGGATCGTGTCCAGAGAGCCATCGTTGACTGACATTTGCGAAGATTTGCTTA 3822
QY 3421 TGGAGACACAGCCGGGTCGTATCAGATGAAGATTTATGACGAGCCAGCCAGGCCAA 3480
Db 3823 TGGAGACACAGCCGGGTCGTATCAGATGAAGATTTATGACGAGCCAGCCAGGCCAA 3882
QY 3481 CATACACCATTCATCAGAGACTCCCTGAGAAATCAACACACAGATGAGACAAAGG 3540
Db 3883 CATACACCATTCATCAGAGACTCCCTGAGAAATCAACACACAGATGAGACAAAGG 3942
QY 3541 AACCAGCTCTGTCGGGCGCAGAACAGCGCATTCGCATAGCTCGCTCTGTTAGCA 3600
Db 3943 AACCAGCTCTGTCGGGCGCAGAACAGCGCATTCGCATAGCTCGCTCTGTTAGCA 4002
QY 3601 GCCTCATATTTTCTTTTGGATGAGATCATCAGCTGCTGATACAGAAAGT 3660
Db 4003 GCCTCATATTTTCTTTTGGATGAGATCATCAGCTGCTGATACAGAAAGT 4062
QY 3661 TGTCCAAAGAGCCCTGAGACAAAGCCAGAGAGAGCCGACCTGATGTCGCCACCG 3720
Db 4063 TGTCCAAAGAGCCCTGAGACAAAGCCAGAGAGAGCCGACCTGATGTCGCCACCG 4122
QY 3721 CTTGTCCACCATTCAGAAATGAGATTAATGATGTTGTTTCAAGATGCAAGTCAAGA 3780
Db 4123 CTTGTCCACCATTCAGAAATGAGATTAATGATGTTGTTTCAAGATGCAAGTCAAGA 4182
QY 3781 GCATGCGACACATCAACAGCTGCTGCGCCAGAAAGCATTTTTCATGCTCAGTGT 3840
Db 4183 GCATGCGACACATCAACAGCTGCTGCGCCAGAAAGCATTTTTCATGCTCAGTGT 4242
QY 3841 CCAGGCTGGAGAGAGCCCTGATGAGCTGATGAGCTGATGAGCTGATGAGCTGAT 3900
Db 4243 CCAGGCTGGAGAGAGCCCTGATGAGCTGATGAGCTGATGAGCTGATGAGCTGAT 4302
QY 3901 ATTGTGTTAAACATGAGCTGATTAATCAAGTTAAAGAGGAGGAGCTGATGAGAAACT 3960
Db 4303 ATT--TGTGTGATATGACATTTTATCAAGTTAA--AGCAACAGCTTACGAAAT 4355
QY 3961 ATGTAAGACTACCTGTTTAACATTTCTGCTGCACTGAAATCATTCACCAAGTTGAG 4020
Db 4356 ATGTAAGAGTATCTGTTTAACATTTCTGCTGCACTGAAATCATTCACCAAGTTGAG 4403
QY 4021 AGTCTCAGATTTTATTAATTAAGAACCA--AAGAACTATTCATGATGAATTAAT 4078
Db 4404 ----AGAGACTTCGTAAATTAAGAACAGAGTGAAGACATCAAGTGAAGAAAT 4458
QY 4079 ACTGCTGTTAATGCAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 4138
Db 4459 CATAGTTTAACATGCAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 4518
QY 4139 TGTATTAATTTTGTATTAATTTT-----ATTGTAAGTACTGCTGCTGAAAGAT 4190
Db 4519 ATGTGTAATTTTGTATTAATTTTCCCATTTTGAGACTGATGACTGCTGCTGAAAGAT 4578
QY 4191 TATAGAAGTGTAAAGACTG--AATGTTGAATTAAGTCTGATATTAATTAATTAAT 4249
Db 4579 TATAGAAGTGTAAAGAAATGATTAATGATTAATGATTAATGATTAATGATTAAT 4635
QY 4250 ACTTTTATAT 4259
Db 4636 ACTTTCATGT 4645

```

RESULT 14

ABL68880
ID ABL68880 standard; DNA; 4646 BP.

XX ABL68880;
AC
XX
XX

DT 15-MAY-2002 (first entry)
XX
XX

DE Kidney cancer related gene sequence SPO ID NO:7217.

```

XX Human; Cancer; colon; breast; ovary; oesophagus; kidney; thyroid;
KW stomach; lung; prostate; pancreas; carcinoma; antitumour;
KW cytostatic; gene therapy; antineoplastic; Wilm's tumour; adenocarcinoma;
XX gene; ds.
OS Homo sapiens.
XX
XX WO200194629-A2.
XX
XX 13-DEC-2001.
XX
XX 30-MAY-2001; 2001WO-US10838.
XX
PR 05-JUN-2000; 2000US-209473P.
PR 05-JUN-2000; 2000US-209531P.
PR 18-SEP-2000; 2000US-233133P.
PR 18-SEP-2000; 2000US-233617P.
PR 20-SEP-2000; 2000US-234009P.
PR 20-SEP-2000; 2000US-234034P.
PR 22-SEP-2000; 2000US-234052P.
PR 22-SEP-2000; 2000US-234509P.
PR 25-SEP-2000; 2000US-234567P.
PR 25-SEP-2000; 2000US-234923P.
PR 25-SEP-2000; 2000US-234924P.
PR 25-SEP-2000; 2000US-235077P.
PR 25-SEP-2000; 2000US-235082P.
PR 25-SEP-2000; 2000US-235134P.
PR 25-SEP-2000; 2000US-235280P.
PR 26-SEP-2000; 2000US-235337P.
PR 27-SEP-2000; 2000US-235638P.
PR 27-SEP-2000; 2000US-235711P.
PR 27-SEP-2000; 2000US-235720P.
PR 27-SEP-2000; 2000US-235840P.
PR 28-SEP-2000; 2000US-235863P.
PR 28-SEP-2000; 2000US-236028P.
PR 28-SEP-2000; 2000US-236032P.
PR 28-SEP-2000; 2000US-236033P.
PR 28-SEP-2000; 2000US-236034P.
PR 28-SEP-2000; 2000US-236109P.
PR 28-SEP-2000; 2000US-236111P.
PR 29-SEP-2000; 2000US-236842P.
PR 29-SEP-2000; 2000US-236891P.
PR 02-OCT-2000; 2000US-237172P.
PR 02-OCT-2000; 2000US-237173P.
PR 02-OCT-2000; 2000US-237278P.
PR 02-OCT-2000; 2000US-237294P.
PR 02-OCT-2000; 2000US-237316P.
PR 02-OCT-2000; 2000US-237316P.
PR 03-OCT-2000; 2000US-237425P.
PR 03-OCT-2000; 2000US-237588P.
PR 03-OCT-2000; 2000US-237604P.
PR 03-OCT-2000; 2000US-237606P.
PR 03-OCT-2000; 2000US-237608P.
PR 01-NOV-2000; 2000US-244867P.
PR 01-NOV-2000; 2000US-245084P.
XX
XX (AVAL-) AVALON PHARM.
XX
XX Young PE, Augustus M, Carter KC, Ebner R, Endress G, Horrigan S;
XX Soppe DR, Weaver Z;
XX WPI; 2002-188264/24.
XX
XX Screening for anti-neoplastic agent involves exposing cells to a
XX chemical agent to be tested for anti-neoplastic activity, and
XX determining a change in expression of a gene of a signature gene set
XX
XX Claim 1; SEQ ID 7217; 44pp; English.
XX
XX The present invention describes a method (M1) for screening for an
XX anti-neoplastic agent. The method involves exposing cells to a chemical
XX agent to be tested for anti-neoplastic activity, determining a change in

```

expression of at least one gene (1) of a signature gene set, where (1) comprises a sequence (S) selected from 8447 sequences (given in ABU61664 to ABU70110), or is at least 95% identical to (S), where a change in expression is indicative of anti-neoplastic activity. (1) has cytostatic activity and can be used in gene therapy. M1 can be used for screening an anti-neoplastic agent, and can be used for producing a product which is the data collected with respect to the anti-neoplastic agent as a result of M1, and the data is sufficient to convey the chemical structure and/or properties of the agent. M1 can be used in the treatment of cancer such as colon, breast, stomach, lung, thyroid, oesophageal, ovarian, kidney, prostate or pancreatic cancer, adenocarcinoma, carcinoma, clear cell cancer, infiltrating ductal cancer, infiltrating lobular cancer, squamous cell carcinoma, neuroendocrine carcinoma, papillary carcinoma and Wilms' tumour.

Sequence 4646 BP; 1371 A; 892 C; 1129 G; 1254 T; 0 other;

Query Match 77.7%; Score 3324.4; DB 24; Length 4646;

Best Local Similarity 87.6%; Pred. No. 0; Mismatches 486; Indels 44; Gaps 8;

Matches 3740; Conservative 0;

1 GGAGCGCGAGTGGGATGATCTGGAAGAGCGCGTGAAGGAGTGCAGAGAAACTT 60

409 GGAGCGCGAGTGGGATGATCTGGAAGAGCGCGTGAAGGAGTGCAGAGAAACTT 60

61 CTGGAAGAGGCAAAAAAGTAAAAAGTGAAGAAAGAAAGAAAGCAACTGTCAG 120

469 CTTTAAAACTGACAAATTAAGTAAAAAGTGAAGAAAGAAAGAAAGCAACTGTCAG 528

121 CAGCTTGCATGTTTGGTATTCATTAATGCTTGAAGTGTATATGTTGGGGGAC 180

529 TGTATTTTCATGTTTGGTATTCATTAATGCTTGAAGTGTATATGTTGGGGGAC 588

181 AATGCGTCATCATTCATGAGTGCACCTCCCTCATGATGCTGTTTGGAAACAT 240

589 TTTGCGCCATCATTCATGAGTGCACCTCCCTCATGATGCTGTTTGGAAACAT 648

241 GAGCATGAGCTTTCATTAATGCTTGAAGTGTATATGTTGGGGGAC 300

649 GACGATATCTTTCATTAATGCTTGAAGTGTATATGTTGGGGGAC 702

301 TGAAGTATTCAGCAATACACAACTTTCATTAATGCTTGAAGTGTATATGTTGGGGGAC 360

703 TAAATGAGTATTCAGCAATACACAACTTTCATTAATGCTTGAAGTGTATATGTTGGGGGAC 762

361 GATGCGTATTCAGCAATACACAACTTTCATTAATGCTTGAAGTGTATATGTTGGGGGAC 822

763 GATGCGTATTCAGCAATACACAACTTTCATTAATGCTTGAAGTGTATATGTTGGGGGAC 882

421 TTTATTCGTTGCTGCGAGCAGCAAGCAATGCTTGAAGTGTATATGTTGGGGGAC 480

823 TTTATTCGTTGCTGCGAGCAGCAGCAATGCTTGAAGTGTATATGTTGGGGGAC 882

481 TGTATTCGTTGCTGCGAGCAGCAGCAATGCTTGAAGTGTATATGTTGGGGGAC 540

883 TGTATTCGTTGCTGCGAGCAGCAGCAATGCTTGAAGTGTATATGTTGGGGGAC 942

541 CCGGCTCAGAGAGATGCTCCAAATCAATGAAGAAATGGGCAAAATTTGGAATGTT 600

943 CCGGCTCAGAGAGATGCTCCAAATCAATGAAGAAATGGGCAAAATTTGGAATGTT 1002

601 CTTTCACTCAATGAAGAAATTTTTCACGGTTTATGAGGTTTACAGTGGTTGGA 660

1003 CTTTCACTCAATGAAGAAATTTTTCACGGTTTATGAGGTTTACAGTGGTTGGA 1062

661 GCTAACCTTGTGATTTGGCCATCAGCCCTTCTTGAGACTTTCAGCCCTGAGG 720

1063 GCTAACCTTGTGATTTGGCCATCAGCCCTTCTTGAGACTTTCAGCCCTGAGG 1122

721 AAGATCTATCTTCAATTCATTAAGAAAGCTTGGCTTATGCAAAAGCTGAGAGCT 780

1123 AAGATCTATCTTCAATTCATTAAGAAAGCTTGGCTTATGCAAAAGCTGAGAGCT 1182

QY 781 AGCTGAAGAAGCTTTAGCAGCAATCAGAACTGATTCGCTTTGGAGCAAAAGAAAGA 840

1183 AGCTGAAGAAGCTTTAGCAGCAATCAGAACTGATTCGCTTTGGAGCAAAAGAAAGA 1242

QY 841 ACTTGAAGGTCACCAAAATTTGAAGAAGCTTAAGAAATTTGGATTAAGAAAGCTAT 900

1243 ACTTGAAGGTCACCAAAATTTGAAGAAGCTTAAGAAATTTGGATTAAGAAAGCTAT 1302

QY 901 CAGGCGCAAACTTTGATTTGTCGCTTTCTTATGATGATGATGATGATGATGATGAT 960

1303 TACAGCCAAATTTGATTTGTCGCTTTCTTATGATGATGATGATGATGATGATGAT 1362

QY 961 TTTCTGATGAGGACCTTGGTCTCTCCAGTGAATATTTGAGACACTACTAC 1020

1363 CTTCTGATGAGGACCTTGGTCTCTCCAGTGAATATTTGAGACACTACTAC 1422

QY 1021 TGTCTTCTTCTGATTTAATTTGGGCTTTTATGATGATGATGATGATGATGATGAT 1080

1423 TGTATCTTCTGATTTAATTTGGGCTTTTATGATGATGATGATGATGATGATGAT 1482

QY 1081 AGCATTTGCAAG 1140

1483 AGCATTTGCAAG 1542

QY 1141 AAGCATTTGCAAG 1200

1543 AAGCATTTGCAAG 1602

QY 1201 ATTCAAAATTTCTGATTTAATTTGGGCTTTTATGATGATGATGATGATGATGAT 1260

1603 ATTCAAAATTTCTGATTTAATTTGGGCTTTTATGATGATGATGATGATGATGAT 1662

QY 1261 CAACCTGAAGTTGAG 1320

1663 CAACCTGAAGTTGAG 1722

QY 1321 GAGCAG 1380

1723 GAGCAG 1782

QY 1381 TGATGAG 1440

1783 TGATGAG 1842

QY 1441 GAGTCAAG 1500

1843 GAGTCAAG 1902

QY 1501 AAATGTCACCAATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1560

1903 AAATGTCACCAATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1962

QY 1561 CATGAAG 1620

1963 CATGAAG 2022

QY 1621 TGGCAG 1680

2023 TGGCAG 2082

QY 1681 GCTGATGAG 1740

2083 GCTGATGAG 2142

QY 1741 GGATTAAG 1800

2143 GGATTAAG 2202

QY 1801 TAATGCCATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1860

2203 TAATGCCATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 2262

QY 1861 TGAATCTATGAG 1920

Db 2263 TGAAGTCAATGAAAGGAGCATTCTACTTAACCTTGTCACATGAGACACAGAGAAA 2322
QY 1921 TGAATTTGATGTAAGAAAATGCGCACTGGTGAATCCAAAAAGTAAAGATGATGCTTGGAAAT 1980
Db 2323 TGAATTTAATTAAGAAAATGACGTGATGTAATCCAAAAAGTAAATGATGCTTGGAAAT 2382
QY 1991 GTCTCCAAAAGATTCAGGGTCCAGTTTAATAAAAAGAGTCAATCCGACAGATATACA 2040
Db 2383 GTCTTCAAAATGATTCAGATCCAGTCTAATTAACAAAAAGTCAATCTGATGAGAGTGTCCG 2442
QY 2041 TGCACCAAGGCCACAGACAGAAAGCTTGGTGAACAAAAAGAGACTTGAATGAGATGTACC 2100
Db 2443 TGGATCCAAAGCCCCAGACAGAAAAGCTTAGTACCAAAAGAGGCTCTGATGAAAGATATCC 2502
QY 2101 TCCAGTTTCCCTTGAGAGATTCCTGAAGCTCAACATCAATGAAATGATTTGTGT 2160
Db 2503 TCCAGTTTCCCTTGGAGAGATTAATGAAGCTAAATTTAACTGAATGGCTTATTTGTGT 2562
QY 2161 TGGATATTTTGTGCTATTATTAACGAGGCGCTGCACACAGCATTTCAATATATTTTC 2220
Db 2563 TGGATATTTTGTGCTATTATTAATGAGGCGCTGCACACAGCATTTGCAATATATTTTC 2622
QY 2221 AAGGATTAATAGGAGTCTTTACCCGAGATGAGATCCTGAACAAACAGACAAATAGTAA 2280
Db 2623 AAGGATTAATAGGAGTCTTTACAAAGATGATATCTGAACAAACAGACAAATAGTAA 2682
QY 2281 CATGTTTCTGATGTTTCTAGTCTTGAATTAATTTCTTTTATTTATTTTCCCTCA 2340
Db 2683 CTGTTTCTGATGTTTCTAGTCTTGAATTAATTTCTTTTATTTTATTTTCCCTCA 2742
QY 2341 GGGCTTCACATTTGGCAAAAGCTGGGAGATCCTCAATAGCGGCTGATACATGTTTT 2400
Db 2743 GGGTTTACATTTGGCAAAAGCTGGGAGATCCTCAACAGGCGCTCGATATACATGTTTT 2802
QY 2401 CAGATCCATGCTGAGACAGAGATGTCAGCTGTTGATGACCTTAAACAAACCCATGAGC 2460
Db 2803 CAGATCCATGCTGAGACAGAGATGTCAGCTGTTGATGACCTTAAACAAACCCATGAGC 2862
QY 2461 ATTGACACCGAGGCTTGGCATGATGCGGCTCAAGTTAAAGGGCTATAGTTCCAGGCT 2520
Db 2863 ATTGACACCGAGGCTTGGCATGATGCGGCTCAAGTTAAAGGGCTATAGTTCCAGGCT 2922
QY 2521 TGTGTCATTAACCCAGATATAGCAATCTTGGACAGGAGTATATATATCTTAATCTA 2580
Db 2923 TGTGTCATTAACCCAGATATAGCAATCTTGGACAGGAGTATATATATCTTAATCTA 2982
QY 2581 TGTGTCATTAACCATTTTACTCTTGAATTTGATACCATCATTTGCAATAGCAGAGT 2640
Db 2983 TGTGTCATTAACCATTTTACTCTTGAATTTGATACCATCATTTGCAATAGCAGAGT 3042
QY 2641 TGTGTAATGAAATGTTGCTGAGACAGCATGAAAGATTAAGAAAGAGCTTGAAGAGC 2700
Db 3043 TGTGTAATGAAATGTTGCTGAGACAGCATGAAAGATTAAGAAAGAGCTTGAAGAGC 3102
QY 2701 TGGGAATATGTCACAGAGCCATGCAAAATCTCGAATGTTGTTCTTACTCGGGA 2760
Db 3103 TGGGAATATGTCACAGAGCCATGCAAAATCTCGAATGTTGTTCTTACTCGGGA 3162
QY 2761 GCAGAAGTTGATATCATATGACAGAGTTTGAAGTACATACAGAAATCTTTGAG 2820
Db 3163 GCAGAAGTTTGAACATATGATGTCAGAGTTTGAAGTACATACAGAAATCTTTGAG 3222
QY 2821 GAAGACACATCTTGGGGTCTCATTTTCTATCACCCAGAGCATGATATTTTCTTA 2880
Db 3223 GAAGACACATCTTGGAGATTTGATATTTCCCTCACCCAGAGCATGATATTTTCTTA 3282
QY 2881 TGTGTCATTTCCGTTTGGTGTACTGTTGCAATATGATTTCAATGATTTTCAAGA 2940
Db 3283 TGTGTCATTTCCGTTTGGAGCTTACTTGTGTGCAATATGATTTTCAAGA 3342
QY 2941 TGTCTTTTGGATTTCTAGCATTTGTCTTGTGTCATGAGTGGGAGGCTCAGTTTC 3000

Db 3343 TGTCTGTAGTATTTTACAGCTGTTGTCTTTGTGTCATGAGCCGCTGGGCACTCAGTTTC 3402
QY 3001 ATTGTCCTGACATATGCAAAAGCCAAAGTATGACAGCCAGCTGATCATGATATCTTA 3060
Db 3403 ATTGTCCTGACATATGCAAAAGCCAAAGTATGACAGCCAGCTGATCATGATATCTTA 3462
QY 3061 AAAAGCCCTGATTTACAGCTTACAGCCCTCAGCGCTCAGCCAAATACGTTGGAG 3120
Db 3463 AAAAGCCCTTGTATTTACAGCTTACAGCGACGAGAGGCTTAATGCGCAACACATTTGGAAG 3522
QY 3121 AAATGTACATTTAATGAGTGTGTTCAACTATCCACTGACACAGCATCCCTGTGT 3180
Db 3523 AAATGTACATTTGTTGAAGTGTATTTCAACTATCCACTGACACAGCATCCCTGTGT 3582
QY 3181 CCAGGGCTGAGGCTGAGGTGTAAGAGGCCAGACGCTGCGCTGTAGTATGACGTGG 3240
Db 3583 TCAGGGACTGAGGCTGAGGTGTAAGAGGCCAGACGCTGCGCTGTAGTATGACGTGG 3642
QY 3241 CTGTGGGAAGACAGATTTGTCAGCTCCTAGAGGCTTATGACCCCTGGCTGGTTC 3300
Db 3643 CTGTGGGAAGACAGATTTGTCAGCTCCTAGAGGCTTATGACCCCTGGCTGGTTC 3702
QY 3301 AGTGTATTTGATGCAAAAGATTAAGCACCTGAATGTCCAGTGGCTCCGACACACT 3360
Db 3703 AGTGTGTGATGCAAAAGATTAAGCACCTGAATGTCCAGTGGCTCCGACACACT 3762
QY 3361 GGGACGCTGCTCAGAGAGCCCATCTGTTGACGTGACGACATTTGCCAGAACTTGCCTA 3420
Db 3763 GGGACGCTGCTCAGAGAGCCCATCTGTTGACGTGACGACATTTGCCAGAACTTGCCTA 3822
QY 3421 TGGAGCAACAGCGGGTGTATCATGATGAAGATTAATGACGAGCCAGAGGAGCAA 3480
Db 3823 TGGAGCAACAGCGGGTGTATCATGATGAAGATTAATGACGAGCCAGAGGAGCAA 3882
QY 3481 CATACACCACTTCAATCGAGACATCCCTGAGAAATTAACAACAGAGTAGAGACAAAG 3540
Db 3883 CATACACCTTCAATCGAGTACTGCTTAATTAATTAATGACAAAGTAGAGACAAAG 3942
QY 3541 AACCCAGCTCTGCTGGGCAAGAAACAGCCATTTGCCATGCTGCGGCTGTTTGAACA 3600
Db 3943 AACCCAGCTCTGCTGGGCAAGAAACAGCCATTTGCCATGCTGCGGCTGTTTGAACA 4002
QY 3601 GCTCATATTTTGGCTTTGATGAGTACATACAGCTCTGATACAGAAAGTAAAGT 3660
Db 4003 GCTCATATTTTGGCTTTGATGAGTACATACAGCTCTGATACAGAAAGTAAAGT 4062
QY 3661 TGTCCAAAGAGCCCTGGAACAAGCCAGAGAGCCAGATGCTGATGTCACCG 3720
Db 4063 TGTCCAAAGAGCCCTGGAACAAGCCAGAGAGCCAGATGCTGATGTCACCG 4122
QY 3721 CTGTCCACCATCCAGATGCAATTTAATAGTGTGTTTCAAGATGCGAAAGTCAAGA 3780
Db 4123 CTGTCCACCATCCAGATGCAATTTAATAGTGTGTTTCAAGATGCGAGTCAAGA 4182
QY 3781 GCATGACACATCAACAGCTGTGCGCCAGAAAGCATCTATTTTCCATGCTAGCT 3840
Db 4183 GCATGACACATCAACAGCTGTGCGCCAGAAAGCATCTATTTTCCATGCTAGCT 4242
QY 3841 CCAGGCTGGAAGAGCCGTAGTGAACCTGTGCCATATGAGCTGTTAATATTTTAAAT 3900
Db 4243 CCAGGCTGGAAGAGCCGTAGTGAACCTGTGACCTATGAGATGTTAATATCTTTAAT 4302
QY 3901 ATTGTGTTAAACATGAGCATTTAATCAAAAGTTAAAGGTAGCAGCTTACGAAAACT 3960
Db 4303 ATT--TGTATGATATGACATTTATTAAGTTAA--AGCAAACTTACGAAAT 4355
QY 3961 ATGTGAATATCTTTTAACTTTCTTGTGCTGCACTGAAAGATTCACCAAGTTGAG 4020
Db 4356 ATGTGAATATCTTTTAACTTTCTTGTGCTGCACTGAAAGTTCAGAGCTTC----- 4403
QY 4021 AGCTTCAGATTTTAAATTAAGAGCAAA--AAGAAACATTAATGATGAAATAAAT 4078
Db 4404 -----AGAGACTGTGTATTAAGAGAACAGAGTGAAGACATCTCAAGTGGAGAAAT 4458

	Query Match	77.6%;	Score 3122.6;	DB 22;	Length 4349;	
	Best Local Similarity 87.5%;	Pred. No. 0;				
	Matches 3740;	Conservative 0;	Mismatches 489;	Indels 44;	Gaps	
QY	1 GGAGCGGAGTGGGATGCTCTGGAAGGAGCGGTAAGGAGAGTGCAGAGAACTT 60					
Db	110 GGAGCGGAGTGGGATGCTCTGGAAGGAGCGGTAAGGAGAGTGCAGAGAACTT 60					
QY	61 CTGGAATATGGGCAAAAAAAGTAAAAAAATGAGAGAAAGAAAGAAACCACTCTGAG 120					
Db	170 CTTTCTTTAACTGAACATATAAGCTGAAAAAGATAAAGAAAGAAAGAAACCACTCTGAG 229					
QY	121 CAGCTTGCATGTTTCCCTATTGCAAAATGGCTGTAGTGTGATATGTTGGTGGAGC 180					
Db	230 TGTATTTCAAGTTCCTGCTATTCAAATGGCTGTAGTGTGATATGTTGGTGGAGC 289					
QY	181 AATGGCTGCATCATCCATGAGAGTGCACCTCCCTCATGATGCTGTTTGGAAACAT 240					
Db	290 TTTGGCTGCATCATCCATGAGAGTGCACCTCCCTCATGATGCTGTTTGGAGAAAT 349					
QY	241 GACAGATACCTTTCGAATGCGAATTCGAGAAATTCGAGAAACCAAACTTTCAGTAAATTA 300					
Db	350 GACAGATACCTTTCGAATGCGAATTCGAGAAATTCGAGAAATTCGAGAAACCAAACTTTCAGTAAATTA 403					
QY	301 TGAAGTATTACAGCAATACCAATTCATCAACCATCTGAGAGAGAAATGACCAC 360					
Db	404 TAAATAGAGATATCAATGATACAGGGTTCCTCATGATCTGGAGAGAACTATGACAG 463					
QY	361 GTATGCCATTAATTACAGTGGGATGCGTGCCTGGCGGCGGCGTGCCTTACATCAAGT 420					
Db	464 GTATGCCATTAATTACAGTGGGATGCGTGCCTGGCGGCGGCGTGCCTTACATCAAGT 523					
QY	421 TTTCTTTGGTGGCTGGCAGACAGAAAGACATCTCAAAATTAAGAAACCAATTTTTTCA 480					
Db	524 TTTCTTTGGTGGCTGGCAGACAGAAATTCACAAATTAAGAAACCAATTTTTTCA 583					
QY	481 TGTATATGCGACAGAGATGGCTGGTTCAGCTGCATGACGTTGGGAGCTTAACAC 540					
Db	584 TGTATATGCGACAGAGATGGCTGGTTCAGCTGCATGACGTTGGGAGCTTAACAC 643					
QY	541 CCGGCTGCACAGACATGCTCTCAAAATCAATGAAGAAATGGCGACAAAATGGGATGTT 600					
Db	644 CCGGCTGCACAGACATGCTCTCAAAATCAATGAAGAAATGGGAGACAAAATGGGATGTT 703					
QY	601 CTTTCATCATTAACCAATTTTTCAACGGTTTTATAGTGGGTTTTACAGTGGTTGGA 660					
Db	704 CTTTCATCATTAACCAATTTTTCAACGGTTTTATAGTGGGTTTTACAGTGGTTGGA 763					
QY	661 GCTAACCTTGTGATTTTGGCCATCAGCCCTGTTCTTGACCTTTCACGCGGCATCTGGGC 720					
Db	764 GCTAACCTTGTGATTTTGGCCATCAGCCCTGTTCTTGACCTTTCACGCGGCATCTGTGGGC 823					
QY	721 AAAGTACTATCTCATTTACTGATTAAGAAACTCTTGAGCTATGCAAAACCTGGAGCAGT 780					
Db	824 AAAGTACTATCTCATTTACTGATTAAGAAACTCTTGAGCTATGCAAAACCTGGAGCAGT 883					
QY	781 AGCTGAGAGAGCTTAGCAGCAATCAGACTGTGATTTGCTTTGGAGAGCAAAAAGAGA 840					
Db	884 AGCTGAGAGAGCTTAGCAGCAATCAGACTGTGATTTGCTTTGGAGAGCAAAAAGAGA 943					
QY	841 ACTTGAAGGATCAACAAAAATTTGAAGAAGCTTAAAGAAATTTGGATTAAGAAAGCTAT 900					
Db	944 ACTTGAAGGATCAACAAAAATTTGAAGAAGCTTAAAGAAATTTGGATTAAGAAAGCTAT 1003					
QY	901 CAGGCGCAACATTTCTATGTTGGCCGCTTCTTATGATCTATGATCATATGCTCTGGC 960					
Db	1004 TAAAGCCATATTTCTATGTTGGCCGCTTCTTATGATCTATGATCATATGCTCTGGC 1063					
QY	961 TTTCTGATGAGACTCTTGTGCTCTCCAGCTGAATTTCTATTTGACAAGTACTCAC 1020					
Db	1064 CTTCTGATGAGACTCTTGTGCTCTCCAGCTGAATTTCTATTTGACAAGTACTCAC 1123					

QY	1021	TGCTCTCTTTCGTGATTAATTGGGGCTTTT	AGTATGAGACGGCATCCCAACATTGA	1080
Db	1124	TGATTCCTTTTCGTATTAATTGGGGCTTTT	AGCTTGGACAGGCATCTCCAACTTGA	1183
QY	1081	AGCATTTGCAACGGCAAGGGAGGACGCTTA	ATGAATCTTCAAGATTAATTTGCATTAACC	1140
Db	1184	AGCATTTGCAAAATGCAAGAGGAGGACGCTTA	ATGAATCTTCAAGATTAATTTGAATTAAGCC	1243
QY	1141	AAGCATTTGACAGCTATTTCCAAAGCTGGACAT	TAACCAATATATTAAGGCAATTTTGA	1200
Db	1244	AAGATTTGACAGCTATTTCCAAAGCTGGACAA	TAACCAATATATTAAGGCAATTTTGA	1303
QY	1201	ATTTCAAAATATGTTTCACTTCAGTTACCTTTC	GTGAAAAGAAAGTTAAGATCTTTAAAGGCT	1260
Db	1304	ATTTCAGAAATGTTTCACTTCAGTTACCTTCG	TAACCAATATATTAAGGCAATTTTGA	1363
QY	1261	CAACCTTAAGGTTTCAAGTGGGACAGCATGTGG	CGCTGGTGGGAACATGGGCTCGGGGAA	1320
Db	1364	GAACCTTAAGGTTTCAAGTGGGACAGCGGTGG	CGCTGGTGGGAACATGGGCTCGGGGAA	1423
QY	1321	GAGCAGACACGCTGACGCTGATGAGAGGCTCT	ATGACCCAGATGGCAATGCTGTAT	1380
Db	1424	GAGCAGACACAGTCCAGCTGATGAGAGGCTCT	ATGACCCAGATGGGGAATGCTGATGCT	1483
QY	1381	TGATGACAGGACATTTAGACCATTAATATPAA	ATGAGCATCTTGGGAAATTAAGCTGTGCT	1440
Db	1484	TGATGACAGGATATTTAGACCATTAATATPAA	ATGAGCATCTTGGGAAATTAAGCTGTGCT	1543
QY	1441	GAGTCAGAGGCTGTGTTTGTTCGCAACAGATG	AGTGTGAACCATTTGCTATGGCCGGGA	1500
Db	1544	GAGTCAGAGGACCTGTATTTGTTTCCACACAG	ATGATGTAACCATTTGCTATGGCCGGGA	1603
QY	1501	AAATGTACCATGTGATGATTTAGAAAGCTGTT	TAAGAAAGCCATGCTATGATTTTAT	1560
Db	1604	AAATGTACCATGTGATGATTTAGAAAGCTGTT	TAAGAAAGCCATGCTATGATTTTAT	1663
QY	1561	CATGAAAACCTACTTAATTAATTTGACACTCT	GTTGGAGAGAGAGGGCCCACTGATGCTG	1620
Db	1664	CATGAAAACCTGCTTCTAATTAATTTGACACT	CCCTGTTGGAGAGAGGGGCCCACTGATGCTG	1723
QY	1621	TGSGACAGAAACAGATGGCCATCTGCTGGGG	CCCTGTTGGCAACCCCAAGATTTCTCT	1680
Db	1724	TGSGACAGAAAGAGAGGATGGCCATCTGCTGG	GGCCCTGTTGGCAACCCCAAGATTTCTCT	1783
QY	1681	GCTGGATGAGGCAACGTCACACTCTGAGACACT	GATGAAAGAGAGGTTCAAGGTGAGTGCCCT	1740
Db	1784	GCTGGATGAGGCCCACTGTCACACTCTGAGACA	CAAGAGAGGTTCAAGGTGAGTGAGTGCCCT	1843
QY	1741	GGATTAAGGCCAGAAAAGGCCGAGCACTACCAT	TGTGATAGCTCATGCTTTGCTACAGTTTCG	1800
Db	1844	GGATTAAGGCCAGAAAAGGTTGGGACCAACCAT	TGTGATAGCTCATGCTTTGCTACAGTTTCG	1903
QY	1801	TAAATGCCAGTCTCTCTGCTGTTTATATATGAG	ATGATGAGGCAATTTGGGAAAGAAATCATGA	1860
Db	1904	TAAATGCTGACGCTCTCTCTGCTGTTTATATAT	GAGATGAGGCAATTTGGGAAAGAAATCATGA	1963
QY	1861	TGAAGCTATGAAAAGAGAGGGGCAATTTACTTT	CAAACTTTGTCACAATGTCAAGACAAGAGAAA	1920
Db	1964	TGAAGCTATGAAAAGAGAGGCAATTTACTTTCA	AACTTTGTCACAATGTCAAGACAAGAGAAA	2023
QY	1921	TGAAAATGAGTTAGAAAATGCCACACTGGTGAT	TCACAAAAGTGAAGAGTGAATGCTTGGAAAT	1980
Db	2024	TGAAAATGAGTTAGAAAATGCGATGATGATTC	CAAAAGTGAAGAGTGAATGCTTGGAAAT	2083
QY	1981	GTCCTCAAAAGATTTCAAGGTGCGAGTTTAAT	TAATTAATTAATCAACGCGAGAGATTAACA	2040
Db	2084	GTCCTCAAAATGATTTCAAGGTGCGAGTTTAAT	TAATTAATTAATCAACGCGAGAGATTAACA	2143
QY	2041	TGCACACAAAGGCCCAACAGAAAAGCTTGGT	TACAAAAGAGGCACTGGAATGAGAAATGATAC	2100
Db	2144	TGCATCACAAAGGCCCAACAGAAAAGCTTGGT	TACAAAAGAGGCTGTGATGAGAAATGATAC	2203
QY	2101	TCCAGTTCTCTTCTGAGAGATTTCTGAAGCTG	AACTCAACTGAATGAGGCTTATTTTGTGCT	2160

[illegible]

Db 3284 TCAGGAGCTGAGCCTGGAGGTGAAGAAAGGCCAGACGCTGGCTCTGCTGGGACAGCTGG 3343

QY 3241 CTGTGGAGAGACAGATGTTGTTGCTCCTAGAGCGCTCTATAGACCCCTTGGCTGGTTC 3300

Db 3344 CTGTGGAGAGACAGAGTGGTCCAGCTCTGTGAGCGGTTCTACACACCCTTGGAGGGAA 3403

QY 3301 AGTCTAATTTATGTGCAAAAGATAAAGCACATGATGTCAGTGGCTCCAGCACACT 3360

Db 3404 AGTCTGCTTATGTGCAAAAGAAATAAAGCGACTGAATGTTCAAGTGGCTCCAGCACACT 3463

QY 3361 GGGCATGCTGTCTGAGAGCCATCTGTTTGACTGCAGCATTGCGAGAAACATGGCTA 3420

Db 3464 GGGCATGCTGTCTGAGAGCCATCTGTTTGACTGCAGCATTGCTGAGAACATGGCTA 3523

QY 3421 TGGAGACAACAGCGGGTGTATCATGAAGATTATGCAGGACGCCAAGAGGCCAA 3480

Db 3524 TGGAGACAACAGCGGGGTGTATCATGAAGATGTTGAGGGGACAGCAAGAGGCCAA 3583

QY 3481 CATACACCACTTCATCGAGACTCCCTGAGAAATACACACAGAGTAGAGACAAAG 3540

Db 3584 CATACATGCTTCATCGAGTCACTCCCTAATTAATATAGCATTAAGTAGAGACAAAG 3643

QY 3541 AACCAGCTCTGTGGTGGCCAGAAAGCCGATGGCATGGCCGCTCTGTTAGACA 3600

Db 3644 AACTGAGCTCTGTGGCCAGAAACAGCATGGCATGGCTCTGTTAGACA 3703

QY 3601 GCCTCATATTTTGTGTTGATGAAGCTATCATGCTCTGATACAGAAAGTGAAGGT 3660

Db 3704 GCCTCATATTTTGTGTTGATGAAGCCAGCTGCTGTGATACAGAAAGTGAAGGT 3763

QY 3661 TGTCCAGAAAGCCCTGAGACAAAGCCAGAGAGGCCGACCTGCTATGTGATGCCACCG 3720

Db 3764 TGTCCAGAAAGCCCTGAGACAAAGCCAGAGAGGCCGACCTGCTATGTGATGCCACCG 3823

QY 3721 CTGTGCACCATTCAGAAATGAATTAATAGTGTGTTTCAAGATGGCAAGTCAAGGA 3780

Db 3824 CTTGTCCACCATTCAGAAATGAATTAATAGTGTGTTTCAAGATGGCAGACTCAAGGA 3883

QY 3781 GCATGGCACACATCAAGCTGTGCGCCAGAAAGCATCTATTTTCCATGGTCAAGT 3840

Db 3884 GCATGGCACAGATCAGAGCTGTGCGCCAGAAAGCATCTATTTTCCATGGTCAAGT 3943

QY 3841 CCAGGCTGAGCAAGCGCTAGTGAAGTGTGCCATATGAGCTGTTAAATATTTTAAAT 3900

Db 3944 CCAGGCTGAGCAAGCGCCAGTGAAGTGTGAGTGAATTAATATTTTAAAT 4003

QY 3901 ATTTGTGTAAACATGGCATTTAATCAAGTTAAAGTGAAGTGAAGTGAAGTGAAGT 3960

Db 4004 ATT--TGTGTGAATATGATTTATTCAAAGTTAA--AGCAACACTTACAGAAAT 4056

QY 3961 ATGTGAACACTGCTGTTTAACTTTGCTGCACTGAGATCAATCCACAAAGTTCAG 4020

Db 4057 ATGAAGAGGTATGTTTAACTTTCTCTGCAAGTTCAGAGTCTC----- 4104

QY 4021 AGTCTCAGATTTATTAATTAAGAACCA--AAGAACATTATCTGATGAATAAAT 4078

Db 4105 ----AGAGACTTCGTATTAAGAAACAGAGAGACATCATCAAGTGAAGAAAT 4159

QY 4079 ACTGTGTATTAATGATTAATAAATATAGATTAATCAAGTGAAGTGTGTTTAAAT 4138

Db 4160 CATAGTTTAACTGCATTAATTAATTAACAGAAATTAAGATTAATTTAAAGATAAA 4219

QY 4139 TGTATATTTTGTGTTATTTT-----ATTGTAATCTGCTTGTGTAAGAT 4190

Db 4220 ATGTGTAATTTTGTGTTATTTTCCCATTTGCACTGTAAGTGTGCTAATAAGAT 4279

QY 4191 TATAGAAAGTGTAAAGTACTG-AAATGTTGAATAAGTGTAGCTATATTAATACTAA 4249

Db 4280 TATAGAAAGTGTAAAGTATGAATGTTGATTAAGTG--TCTATATTAATACTAA 4336

QY 4250 ACTTTATATCAA 4262

Db 4337 ACTTCAAAAAA 4349

Search completed: December 9, 2002, 21:27:35
Job time : 1066 secs

GenCore version 5.1.3
Copyright (c) 1993 - 2002 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: December 9, 2002, 17:12:12 ; Search time 866 Seconds
(Without alignments)
1926.700 Million cell updates/sec

Title: US-09-672-725c-1

Perfect score: 4279

Sequence: 1 ggaagcgcgaggtcgagatg.....caaaaaaaaaaaaaaaaaa 4279

Scoring table: IDENTITY NUC

Searched: 350425 seqs, 194966369 residues

Total number of hits satisfying chosen parameters: 700850

Minimum DB seq length: 0

Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : Published Applications_NA.*
1: /cgn2_6/ptodata/1/pubpna/US07_PUBCOMB.seq.*
2: /cgn2_6/ptodata/1/pubpna/PCT_NEW_PUB.seq.*
3: /cgn2_6/ptodata/1/pubpna/US06_NEW_PUB.seq.*
4: /cgn2_6/ptodata/1/pubpna/US06_PUBCOMB.seq.*
5: /cgn2_6/ptodata/1/pubpna/US07_NEW_PUB.seq.*
6: /cgn2_6/ptodata/1/pubpna/US07_PUBCOMB.seq.*
7: /cgn2_6/ptodata/1/pubpna/US08_NEW_PUB.seq.*
8: /cgn2_6/ptodata/1/pubpna/US08_PUBCOMB.seq.*
9: /cgn2_6/ptodata/1/pubpna/US09_NEW_PUB.seq.*
10: /cgn2_6/ptodata/1/pubpna/US09_PUBCOMB.seq.*
11: /cgn2_6/ptodata/1/pubpna/US10_NEW_PUB.seq.*
12: /cgn2_6/ptodata/1/pubpna/US10_PUBCOMB.seq.*
13: /cgn2_6/ptodata/1/pubpna/US60_NEW_PUB.seq.*
14: /cgn2_6/ptodata/1/pubpna/US60_PUBCOMB.seq.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	4221.8	98.7	4317	9	US-10-044-671-1
2	3316.4	77.5	4643	9	Sequence 1, Appl1
3	3226.6	75.4	8630	10	US-09-1072-621-2
4	3223.4	75.3	8630	10	US-09-306-417-2
5	3202	74.8	3860	10	US-09-306-417-1
6	3200.4	74.8	3860	10	US-09-866-866A-1
7	2936.2	68.6	4788	10	US-09-866-866A-3
8	2920.6	68.3	4425	10	US-09-866-866A-7
9	22917	68.2	4369	10	US-09-769-097-3
10	2716.6	62.5	4369	10	US-09-769-097-1
11	2677.4	62.6	4254	10	US-09-866-866A-5
12	2240	52.3	3924	10	US-09-917-800A-1424
13	2081.8	48.7	3912	10	US-09-880-107-2299
14	1323.8	30.9	3621	10	US-09-917-800A-1560
15	1233.8	28.8	3702	10	US-09-873-409-14
16	1130	26.4	3177	10	US-09-873-409-13
17	1065.2	24.9	5010	10	US-09-873-409-12
18	1026.6	24.0	2856	10	US-09-917-800A-483
19	772.2	18.0	4175	10	US-09-873-409-10
					US-09-749-340-3

20	702.4	16.4	2066	10	US-09-873-409-9	Sequence 9, Appl1
21	655.4	15.3	3861	9	US-09-938-842A-2237	Sequence 2237, Ap
22	636.2	14.9	3861	9	US-09-938-842A-263	Sequence 263, Ap
23	621.4	14.5	1940	10	US-09-873-409-16	Sequence 16, Appl
24	606.8	14.2	4002	9	US-09-758-828-1	Sequence 1, Appl1
25	606.8	14.2	4002	9	US-09-758-828-3	Sequence 3, Appl1
26	539.4	12.6	2698	10	US-09-749-340-5	Sequence 5, Appl1
27	530.4	12.4	2021	10	US-09-873-409-15	Sequence 15, Appl
28	495.6	11.6	3512	10	US-09-749-340-7	Sequence 7, Appl1
29	402	9.4	4653	9	US-10-101-388-2	Sequence 2, Appl1
30	399.4	9.3	2681	10	US-09-749-340-8	Sequence 8, Appl1
31	355.2	8.3	1810	10	US-09-749-340-4	Sequence 4, Appl1
32	324.2	7.6	1175	10	US-09-873-409-11	Sequence 11, Appl
33	285.6	6.7	452	10	US-09-728-445-603	Sequence 603, App
34	261.8	6.1	8036	9	US-10-101-388-1	Sequence 1, Appl1
35	218.2	5.1	3512	9	US-10-072-621-1	Sequence 18, Appl
36	216.2	5.1	2298	10	US-09-795-693-18	Sequence 18, Appl
37	216.2	5.1	3408	10	US-09-795-693-16	Sequence 16, Appl
38	206.2	4.8	981	10	US-09-770-445-272	Sequence 272, App
39	205.4	4.8	939	10	US-09-974-300-6511	Sequence 13, Appl
40	203	4.7	1076	9	US-09-764-884-13	Sequence 21, Appl
41	194.8	4.6	1812	10	US-09-815-242-1147	Sequence 6018, Ap
42	192	4.5	1367	9	US-09-764-884-21	Sequence 2176, Ap
43	188.6	4.4	1749	10	US-09-815-242-6018	Sequence 8403, Ap
44	187.4	4.4	1233	10	US-09-974-300-2176	
45	185.2	4.3	1737	10	US-09-815-242-8403	

ALIGNMENTS

```
RESULT 1
US-10-044-671-1
; Sequence 1, Application US/10044671
; Patent No. US20020171747A1
; GENERAL INFORMATION:
; APPLICANT: Washington State University Research Foundation
; APPLICANT: Mealey, Katrina
; APPLICANT: Bentjen, Steven
; TITLE OF INVENTION: MDRI VARIANTS AND METHODS FOR THEIR USE
; FILE REFERENCE: 4630-61733
; CURRENT APPLICATION NUMBER: US/10/044.671
; CURRENT FILING DATE: 2002-01-10
; PRIOR APPLICATION NUMBER: US 60/261,578
; PRIOR FILING DATE: 2001-01-12
; PRIOR APPLICATION NUMBER: US 60/314,829
; PRIOR FILING DATE: 2001-08-24
; NUMBER OF SEQ ID NOS: 10
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 1
; LENGTH: 4317
; TYPE: DNA
; ORGANISM: Canis familiaris
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (70)..(3912)
; OTHER INFORMATION:
US-10-044-671-1
Query Match          98.7%: Score 4221.8; DB 9; Length 4317;
Best Local Similarity 99.5%: Pred. No 0;
Matches 4245; Conservative 0; Mismatches 17; Indels 3; Gaps 1;
QY 3 AGCCGAGGTGCGATGATCTGTAAGAGAGCCGTAAGGGAGTGCAGAGAAACTTC 62
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 56 AGCCGAGGTGCGATGATCTGTAAGAGAGCCGTAAGGGAGTGCAGAGAAACTTC 115
QY 63 GGAATAATGGGCAAAAAAGTAAAAAATGAGAAAGAAAGAAACCACTGTGACGA 122
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 116 GGAATAATGGGCAAAAAAGTAAAAA---GAGAAAGAAAGAAACCACTGTGACGA 172
QY 123 CGTTTGCAATGTTTCGCTATTCAAATTGCGTGTATGTTATGTTGGTGGGACAA 182
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
```

Db	173	CGTTGCAATGTTTCGCTATTCAAATTCGCTTGATAGGTTGTAATGTTGGTGGGACAA	232
Qy	183	TGGCGCATCATCATGATGAGCTGCATCCCTCTCATGATCTGGTTTTTGGAAACATGA	242
Db	223	TGGCTGCCATCTCCATGAGCTGCATCCCTCTCATGATCTGGTTTTTGGAAACATGA	292
Qy	243	CAGATAGCTTGGCAAAATCGAGAAATTTCAAGAAACAAATTTTCCAGTTAATTAATG	302
Db	293	CAGATAGCTTGGCAAAATCGAGAAATTTCAAGAAACAAATTTTCCAGTTAATTAATG	352
Qy	303	AAAGTATTACGAACATACACAACATTTTCATCAACCATCTGGAGGAGAAATGACACAGT	362
Db	353	AAAGTATTACGAACATACACAACATTTTCATCAACCATCTGGAGGAGAAATGACACAGT	412
Qy	363	ATGCCATATTATACAGTGGGATCGGTGCTGCGTGGTGGCTGGTGGCTTACATCCAGTTT	422
Db	413	ATGCCATATTATACAGTGGGATCGGTGCTGCGTGGTGGCTGGTGGCTTACATCCAGTTT	472
Qy	423	CATTCTGTCCTGGCAGCAGAGACAGATCTCAAAAATTTGAAAACAAATTTTTCATG	482
Db	473	CATTCTGTCCTGGCAGCAGAGACAGATCTCAAAAATTTGAAAACAAATTTTTCATG	532
Qy	483	CTATCATGCGACAGAGATTTGGCTGTTGACGTGCATGACGTTGGGAGCTTAACACCC	542
Db	533	CTATCATGCGACAGAGATTTGGCTGTTGACGTGCATGACGTTGGGAGCTTAACACCC	592
Qy	543	GGCTCAGACAGATGTCCTCAAAATCAATGAAGAAATTTGGCACAATAATGATGTTCT	602
Db	593	GGCTCAGACAGATGTCCTCAAAATCAATGAAGAAATTTGGCACAATAATGATGTTCT	652
Qy	603	TTTACATATAGCAACATTTTTCACCGGTTTTATAGTGGGTTTTACGTGGTTGAAGC	662
Db	653	TTTACATATAGCAACATTTTTCACCGGTTTTATAGTGGGTTTTACGTGGTTGAAGC	712
Qy	663	TAAACCTTGATTTGGCCATCAGCCCTGTTCTTGACCTTTCAGCCGCAATCTGGGCAA	722
Db	713	TAAACCTTGATTTGGCCATCAGCCCTGTTCTTGACCTTTCAGCCGCAATCTGGGCAA	772
Qy	723	AGATTAATCTTCAATTTACTGATTAAGAACTCTTGGCCTATGCAAAAAGCTGGAGCAGTAG	782
Db	773	AGATTAATCTTCAATTTACTGATTAAGAACTCTTGGCCTATGCAAAAAGCTGGAGCAGTAG	832
Qy	783	CTGGAAGAGCTTACAGCAATCAGAACTGATTCCTTTGGAGGACAAAAGAAAGAAC	842
Db	833	CTGGAAGAGCTTACAGCAATCAGAACTGATTCCTTTGGAGGACAAAAGAAAGAAC	892
Qy	843	TTGAAGGTACAAACAAAATTTAGAGAGCTTAAGGAATTTGGATTAAGAAAGCTATCA	902
Db	893	TTGAAGGTACAAACAAAATTTAGAGAGCTTAAGGAATTTGGATTAAGAAAGCTATCA	952
Qy	903	CGGCCAATTTCTATTTGGGCGCTTCTTATTTGATCTATGATCATATGCTGGCTT	962
Db	953	CGGCCAATTTCTATTTGGGCGCTTCTTATTTGATCTATGATCATATGCTGGCTT	1012
Qy	963	TCTGTAATGAGACCTCTTGGTCTCTCAAGTAATTTCTATTTGACAAGTACTACTG	1022
Db	1013	TCTGTAATGAGACCTCTTGGTCTCTCAAGTAATTTCTATTTGACAAGTACTACTG	1072
Qy	1023	TCTTCTTTTCTGTAATTAATTTGGGCTTTTATTTGATTTGAGACAGCATCTCCCAAGCATTTAG	1082
Db	1073	TCTTCTTTTCTGTAATTAATTTGGGCTTTTATTTGATTTGAGACAGCATCTCCCAAGCATTTAG	1132
Qy	1083	CATTGCAAGAGCAGAGACAGCTTATGAAATCTTCAAGATTAATTTGACAAATTAACAA	1142
Db	1133	CATTGCAAGAGCAGAGACAGCTTATGAAATCTTCAAGATTAATTTGACAAATTAACAA	1192
Qy	1143	GCATTGACAGCTATTTCGAAGATGAGACATAAACCAGATTAATTTAAGGAAATTTGCAAT	1202
Db	1193	GCATTGACAGCTATTTCGAAGATGAGACATAAACCAGATTAATTTAAGGAAATTTGCAAT	1252
Qy	1203	TCAAAATATGTTCACTTCAGTTACCTTCTCGAAAAGAAAGTAAAGATCTTAAAGGCTCA	1262
Db	1253	TCAAAATATGTTCACTTCAGTTACCTTCTCGAAAAGAAAGTAAAGATCTTAAAGGCTCA	1312
Qy	1263	ACCTGAAGGTTTCAGAGTGGGACAGACAGTGGCCCTGTTGGGAAACAGTGGCTCGGGAGAA	1322
Db	1313	ACCTGAAGGTTTCAGAGTGGGACAGACAGTGGCCCTGTTGGGAAACAGTGGCTCGGGAGAA	1372
Qy	1323	GCAAGACCGTGCAGCTGATGCAGAGGCTCTATGACCCACAGTATGCTATGATG	1382
Db	1373	GCAAGACCGTGCAGCTGATGCAGAGGCTCTATGACCCACAGTATGCTATGATG	1432
Qy	1383	ATGGACAGGACATTAGACCATTAATGTAAGGATCTGCGGAAATTTACTGCTGTGTGA	1442
Db	1433	ATGGACAGGACATTAGACCATTAATGTAAGGATCTGCGGAAATTTACTGCTGTGTGA	1492
Qy	1443	GTCAGAGCCTGTGTTGTTTGGCACACAGATAGCTGAAAACATTCCTATGCGCGGAAA	1502
Db	1493	GTCAGAGCCTGTGTTGTTTGGCACACAGATAGCTGAAAACATTCCTATGCGCGGAAA	1552
Qy	1503	ATGTCACCATGATGAGATTTGACAAAGCTGTTAAGGAAGCCATAGCTATGATTTATCA	1562
Db	1553	ATGTCACCATGATGAGATTTGACAAAGCTGTTAAGGAAGCCATAGCTATGATTTATCA	1612
Qy	1563	TGAACCTAACCTAATTAATTTGACACTCTGTTGGAGAGAGAGGGCCACCTGATGCTG	1622
Db	1613	TGAACCTAACCTAATTAATTTGACACTCTGTTGGAGAGAGAGGGCCACCTGATGCTG	1672
Qy	1623	GACAGAAAACAGAGATGCCATTTGCTGGGCCCTGTTCCGACACCCACATTTCTGCG	1682
Db	1673	GACAGAAAACAGAGATGCCATTTGCTGGGCCCTGTTCCGACACCCACATTTCTGCG	1732
Qy	1683	TGATGAGGCAAGCTGCAGCTGTGACACTGAAATGGAAGCAGTGGTTCAAGTGGCCCTGG	1742
Db	1733	TGATGAGGCAAGCTGCAGCTGTGACACTGAAATGGAAGCAGTGGTTCAAGTGGCCCTGG	1792
Qy	1743	ATAAGGCCAGAAAAGGCCGAGCTACCATTTGATAGCTCATGTTGTTCTACAGTTGCTA	1802
Db	1793	ATAAGGCCAGAAAAGGCCGAGCTACCATTTGATAGCTCATGTTGTTCTACAGTTGCTA	1852
Qy	1803	ATGCCGATGTCAATGCTGCTGTTTGTATGATGATGAGATCATTTGAGAAAGAAATGATG	1862
Db	1853	ATGCCGATGTCAATGCTGCTGTTTGTATGATGATGAGATCATTTGAGAAAGAAATGATG	1912
Qy	1863	AACCTATGAAGAGAGAGGCAATTTACTTCAAACTGTGCACATGCGACAGAGAGAAATG	1922
Db	1913	AACCTATGAAGAGAGAGGCAATTTACTTCAAACTGTGCACATGCGACAGAGAGAAATG	1972
Qy	1923	AAATTGAGTTAGAAAATGCCACTGTGTAATCCAAAAGTGAAGTATGCTTGGAAATGT	1982
Db	1973	AAATTGAGTTAGAAAATGCCACTGTGTAATCCAAAAGTGAAGTATGCTTGGAAATGT	2032
Qy	1983	CTCCAAAAGATTCAGAGGTCAGCTTATTAATAAAGAAAGTCAACTGCGAGGATATACATG	2042
Db	2033	CTCCAAAAGATTCAGAGGTCAGCTTATTAATAAAGAAAGTCAACTGCGAGGATATACATG	2092
Qy	2043	CACCAAGAGCCAGACAGAAAGCTTGTACAAAAGAGAGACTTGATGAGATGTACTC	2102
Db	2093	CACCAAGAGCCAGACAGAAAGCTTGTGTACAAAAGAGAGACTTGATGAGATGTACTC	2152
Qy	2103	CAGTTTCTTCTGAGAGATTTGAAAGCTGAACACTGAAGTAAAGGCTTATTTTGGTTG	2162
Db	2153	CAGTTTCTTCTGAGAGATTTGAAAGCTGAACACTGAAGTAAAGGCTTATTTTGGTTG	2212
Qy	2163	GTAATATTTGTGTATTTAATAACGAGGCGCTCAACACAGCATTTTCAATATATTTTCAA	2222
Db	2213	GTAATATTTGTGTATTTAATAACGAGGCGCTCAACACAGCATTTTCAATATATTTTCAA	2272
Qy	2223	GGATTATAGGATCTTTTACCCGAGATGAGATCTGAAAACAAACGACAGATTAATACA	2282
Db	2273	GGATTATAGGATCTTTTACCCGAGATGAGATCTGAAAACAAACGACAGATTAATACA	2332
Qy	2283	TGTTTTCTGTATTTCTTACAGTCTTGAATTAATTTCTTTTACATTTTCTCCACAG	2342
Db	2333	TGTTTTCTGTATTTCTTACAGTCTTGAATTAATTTCTTTTACATTTTCTCCACAG	2392

QY 2443 GCTTCACATTTGGCAAGCTGGGAGATTCCTACATAGCGGCTTCGATACATGGTTTCA 2402
|||||
Db 2393 GCTTCACATTTGGCAAGCTGGGAGATTCCTACATAGCGGCTTCGATACATGGTTTCA 2452
QY 2403 GATCCATGCTGAGACAGATGTCAGCTGTTTGTATGACCTTAAAAACACACTGAGCAT 2462
|||||
Db 2453 GATCCATGCTGAGACAGATGTCAGCTGTTTGTATGACCTTAAAAACACACTGAGCAT 2512
QY 2463 TGAACACAGGCTTGGCCAAATGATGCGGCTCAAGTTAAAGGGCTTATAGTCCAGGCTTG 2522
|||||
Db 2513 TGAACACAGGCTTGGCCAAATGATGCGGCTCAAGTTAAAGGGCTTATAGTCCAGGCTTG 2572
QY 2523 CTGTCAATTACCCAGATATATAGCAATCTTGGACAGGCAATTATATCTTAAATCATG 2582
|||||
Db 2573 CTGTCAATTACCCAGATATATAGCAATCTTGGACAGGCAATTATATCTTAAATCATG 2632
QY 2583 GTTGGCAATTACATTTTACTTTCATATGACCATTTGACATATGACATAGCAGATGG 2642
|||||
Db 2633 GTTGGCAATTACATTTTACTTTCATATGACCATTTGACATATGACATAGCAGATGG 2692
QY 2643 TTGAAATGAAATGTTGCTGCGAGACACATGAAATATAGAAAGCTAGAGAGCTG 2702
|||||
Db 2693 TTGAAATGAAATGTTGCTGCGAGACACATGAAATATAGAAAGCTAGAGAGCTG 2752
QY 2703 GGAAGATTGCTACAGAACCCATCGAAACTTCCGAATGTTGTTCTTGTACTCGGAGC 2762
|||||
Db 2753 GGAAGATTGCTACAGAACCCATCGAAACTTCCGAATGTTGTTCTTGTACTCGGAGC 2812
QY 2763 AGAAGTTTAATATCATGTATGACAGAGTTTGCAGTATACATACAGAACTTGTGAGA 2822
|||||
Db 2813 AGAAGTTTAATATCATGTATGACAGAGTTTGCAGTATACATACAGAACTTGTGAGA 2872
QY 2823 AAGCAACATCTTGGGGGTCTCATTTTCTATCCAGGCAATGATATTTTCCATG 2882
|||||
Db 2873 AAGCAACATCTTGGGGGTCTCATTTTCTATCCAGGCAATGATATTTTCCATG 2932
QY 2883 CTGGCTGTTCCGATTTGGTGGCTTACTTGGTGGCAATGATTCATGAACTTTCAGATG 2942
|||||
Db 2933 CTGGCTGTTCCGATTTGGTGGCTTACTTGGTGGCAATGATTCATGAACTTTCAGATG 2992
QY 2943 TTCTTTGGTATTTTCACGCTATTTGCTTGGTGGCAATGATTCATGAACTTTCAGATG 3002
|||||
Db 2993 TTCTTTGGTATTTTCACGCTATTTGCTTGGTGGCAATGATTCATGAACTTTCAGATG 3052
QY 3003 TTCTCTCTACATATCCCAAGCAATATACAGACCCAGCTCATGATCATTTTGA 3062
|||||
Db 3053 TTCTCTCTACATATCCCAAGCAATATACAGACCCAGCTCATGATCATTTTGA 3112
QY 3063 AAAGCCCTTGATGACAGCTACAGCCCTCAGCGGCTCAAGCCAAATACGTTGGAGAA 3122
|||||
Db 3113 AAAGCCCTTGATGACAGCTACAGCCCTCAGCGGCTCAAGCCAAATACGTTGGAGAA 3172
QY 3123 ATGTGACATTTAATAGAGTGTGTTCAATATCCCATGACAGACATATCCCGGTGTC 3182
|||||
Db 3173 ATGTGACATTTAATAGAGTGTGTTCAATATCCCATGACAGACATATCCCGGTGTC 3232
QY 3183 AGGGGCTGAGCTCGAGGTGAAGAGGGCCAGAGCTGGCCCTCGTAGATGACAGTGGCT 3242
|||||
Db 3233 AGGGGCTGAGCTCGAGGTGAAGAGGGCCAGAGCTGGCCCTCGTAGATGACAGTGGCT 3292
QY 3243 GTGGGAAGAGCACAGTTGTTAGCTCCTAGAGGCTTCTATAGACCCCTTGAGCTGTTCA 3302
|||||
Db 3293 GTGGGAAGAGCACAGTTGTTAGCTCCTAGAGGCTTCTATAGACCCCTTGAGCTGTTCA 3352
QY 3303 TTGCTAATGATGCAAAAGATTAAGACCTGAATGCGAGGGCTCCGAGACACACTGG 3362
|||||
Db 3353 TTGCTAATGATGCAAAAGATTAAGACCTGAATGCGAGGGCTCCGAGACACACTGG 3412
QY 3363 GCATCTGTTCTAGAGCCCATCTGTTGATGTCAGCATTCGAGAAATCTGATG 3422
|||||
Db 3413 GCATCTGTTCTAGAGCCCATCTGTTGATGTCAGCATTCGAGAAATCTGATG 3472
QY 3423 GAGACAACAGCCGGTCTATCATGATGAAGATTTATGACGACACCAAGAGGSCAACA 3482

Db 3473 GAGACAACAGCCGGTCTATCATGATGAAGATTTATGAGGACCCAAAGAGGCCAACA 3532
|||||
QY 3483 TACACACTTCATGAGACACACTCCCTGAGAAATACAAACAGATAGAGACAAAGAA 3542
|||||
Db 3533 TACACACTTCATGAGACACACTCCCTGAGAAATACAAACAGATAGAGACAAAGAA 3592
QY 3543 CCAGCTCTGAGGCGAGAAAGAGGCAATGAGCATGAGCTCGGCTGTTAGACAGC 3602
|||||
Db 3593 CCAGCTCTGAGGCGAGAAAGAGGCAATGAGCATGAGCTCGGCTGTTAGACAGC 3652
QY 3603 CTCATATTTTGGCTTTGGATGAACTTACATCAGCTCTGGATTCAGAAAGTGAAGGTTG 3662
|||||
Db 3653 CTCATATTTTGGCTTTGGATGAACTTACATCAGCTCTGGATTCAGAAAGTGAAGGTTG 3712
QY 3663 TCCAAAGAGCCCTGAGCAAAAGCCAGAAAGGCGGACCTGGATTTGTATGAGCCACCGCT 3722
|||||
Db 3713 TCCAAAGAGCCCTGAGCAAAAGCCAGAAAGGCGGACCTGGATTTGTATGAGCCACCGCT 3772
QY 3723 TGTCCACCATCCAGAAATGCAAGATTTAATAGTGTGTTTCAGAAATGCGCAAGTCAAGGAGC 3782
|||||
Db 3773 TGTCCACCATCCAGAAATGCAAGATTTAATAGTGTGTTTCAGAAATGCGCAAGTCAAGGAGC 3832
QY 3783 ATGGCACATCAACAGCTGCTGGCCAGAAAGCATCTATTTTTCATGGTCAAGTCC 3842
|||||
Db 3833 ATGGCACATCAACAGCTGCTGGCCAGAAAGCATCTATTTTTCATGGTCAAGTCC 3892
QY 3843 AGGCTGAGGCAAGAGGCTAGTGAACCTGGCCATATGAGCTGTTTAAATATTTTAAATAT 3902
|||||
Db 3893 AGGCTGAGGCAAGAGGCTAGTGAACCTGGCCATATGAGCTGTTTAAATATTTTAAATAT 3952
QY 3903 TTGTGTTAAACATGAGCAATTAATCAAGTTAAAGGAGACACTTCTGAAAACTAT 3962
|||||
Db 3953 TTGTGTTAAACATGAGCAATTAATCAAGTTAAAGGAGTGAAGCTTACTGAAACATAT 4012
QY 3963 GTAGAACTACCTGTTTAAATCTTCTGTCGCAATGAAATCATCTCCACAGTTCCAGAC 4022
|||||
Db 4013 GTAGAACTACCTGTTTAAATCTTCTGTCGCAATGAAATCATCTCCACAGTTCCAGAC 4072
QY 4023 TCTTCAGATTTTAAATTAATTAAGAACCAAAAGCAATATCTGATGGAATTAATAGT 4082
|||||
Db 4073 TCTTCAGATTTTAAATTAATTAAGAACCAAAAGCAATATCTGATGGAATTAATAGT 4132
QY 4083 GGTATATGCAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 4142
|||||
Db 4133 GGTATATGCAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 4192
QY 4143 TAATTTTGTATTAATTTATTTGTAACTTACTGCTTGTGTAAGATTAATTAAGATGCT 4202
|||||
Db 4193 TAATTTTGTATTAATTTATTTGTAACTTACTGCTTGTGTAAGATTAATTAAGATGCT 4252
QY 4203 AAAAGTACTGAATGTTTGAATTAAGTCTAGCTATTAATTAAGTCTTATATCA 4262
|||||
Db 4253 AAAAGTACTGAATGTTTGAATTAAGTCTAGCTATTAATTAAGTCTTATATGAA 4312
QY 4263 AAAAA 4267
|||||
Db 4313 AAAAA 4317

RESULT 2
US-10-072-621-2
; Sequence 2, Application US/10072621
; Patent No. US20020169137A1
; GENERAL INFORMATION:
; APPLICANT: Renner, Peter B.
; APPLICANT: Connop, Bruce P.
; APPLICANT: Pollard, Michelle
; TITLE OF INVENTION: REGULATION OF AMYLOID PRECURSOR PROTEIN EXPRESSION
; FILE REFERENCE: 100103.402
; CURRENT APPLICATION NUMBER: US/10/072,621
; CURRENT FILING DATE: 2002-02-08


```
Db 2380 GTCTCAATGATTCAGATCCAGTCTAATAGAAAAAGATCACTCGTAGAGTGTCCG 2439
Qy 2041 TGCACCAAGGCCAAGACAGAAAAGCTTGGTACAAAAGAGACTGATAGATGATAC 2100
Db 2440 TGGATCCAAAGCCCAAGACAGAAAGCTTAGTACCAAAAGAGGCGTGGATGAAGATATACC 2499
Qy 2101 TCCAGTTCCCTCTGAGAGATTCGANGCTGACACTGAATGAGGCTTATTTTGTGCT 2160
Db 2500 TCCAGTTCCCTCTGAGAGATTCGANGCTGACACTGAATGAGGCTTATTTTGTGCT 2559
Qy 2161 TGGTATATTTTGTCTATTAATAAAGAGGCTGCACACAGATTTCAATATATTTTC 2220
Db 2560 TGGTATATTTTGTCTATTAATAAAGAGGCTGCACACAGATTTCAATATATTTTC 2619
Qy 2221 AAGGATATATAGGATCTTACCCGATGAGATCTGAAACAAAGACAGATATAGTAA 2280
Db 2620 AAGGATATATAGGATCTTACCCGATGAGATCTGAAACAAAGACAGATATAGTAA 2679
Qy 2281 CATGTTTCCTGATGCTTCTAGTCTAGTCTGATGATTTATTTCTTATTAATTTTCTCCA 2340
Db 2680 CTGTTTCTACTATGTTGTTCTAGCCTTGGATTAATTTCTTATTAATTTTCTTCA 2739
Qy 2341 GGGCTTACATTTGGCAAAAGCTGGAGATCTCTACTAAGCGCTTCATACATGTTTT 2400
Db 2740 GGGCTTACATTTGGCAAAAGCTGGAGATCTCTACTAAGCGCTTCATACATGTTTT 2799
Qy 2401 CAGATCCATGCTGAGACAGAGATGTCAGTGTGATGACCCCTAAACACACACTGGAGC 2460
Db 2800 CCGATCCATGCTGAGACAGAGATGTCAGTGTGATGACCCCTAAACACACACTGGAGC 2859
Qy 2461 ATTGACACAGGCTTGGCANTGATGGGCTCAAGTTAAAGGGCTATAGGTTCCAGCT 2520
Db 2860 ATTGACACAGGCTTGGCANTGATGGGCTCAAGTTAAAGGGCTATAGGTTCCAGCT 2919
Qy 2521 TGTCTGATATACCCAGATATAGCAATCTTGGACAGGCTTATATATCTTAACTCA 2580
Db 2920 TGTCTGATATACCCAGATATAGCAATCTTGGACAGGCTTATATATCTTAACTCA 2979
Qy 2581 TGGTGGCAATTAACACTTTTACTCTTAGCAATTTGATACCATGATGCAATAGCAGAGT 2640
Db 2980 TGGTGGCAATTAACACTTTTACTCTTAGCAATTTGATACCATGATGCAATAGCAGAGT 3039
Qy 2641 TGGTGAATATGAATAATGTTGTCTGGACAAGCACTGAAGAATAGAAAGGCTAGAAAGAGC 2700
Db 3040 TGGTGAATATGAATAATGTTGTCTGGACAAGCACTGAAGAATAGAAAGGCTAGAAAGAGC 3099
Qy 2701 TGGGAAGATTTGCTACAGAAAGCATTCGAACTGTTGTTTCTTACTCGGGA 2760
Db 3100 TGGGAAGATTTGCTACAGAAAGCATTCGAACTGTTGTTTCTTACTCGGGA 3159
Qy 2761 GCAGAGTTTGAATACATGTATGCACAGAGTTTGCAGATACATACAGAAACTCTTGAG 2820
Db 3160 GCAGAGTTTGAATACATGTATGCACAGAGTTTGCAGATACATACAGAAACTCTTGAG 3219
Qy 2821 GAAGAAGACATCTTGGGGTCTCATTTTGTATCACCCAGGAGATGATGATTTTCTTA 2880
Db 3220 GAAGAAGACATCTTGGGGTCTCATTTTGTATCACCCAGGAGATGATGATTTTCTTA 3279
Qy 2881 TGTGCGCTGTTCCGTTTGTGCTTACTGTTGCTGAGAAATGAGTTCAATGAACTTCA 2940
Db 3280 TGTGCGCTGTTCCGTTTGTGCTTACTGTTGCTGAGAAATGAGTTCAATGAACTTCA 3339
Qy 2941 TGTTCCTTTGATTTCTAGCTATTTGCTTGTGTCATGAGAGTGGGGAGCTCACTTC 3000
Db 3340 TGTTCCTTTGATTTCTAGCTATTTGCTTGTGTCATGAGAGTGGGGAGCTCACTTC 3399
Qy 3001 ATTGTCCTCGACTATGCAAAAGCCAAAGTATCAGACGGCCAGCTCATCATGATCATGTA 3060
Db 3400 ATTGTCCTCGACTATGCAAAAGCCAAAGTATCAGACGGCCAGCTCATCATGATCATGTA 3459
Qy 3061 AAAAAGCCCTTGATTTGACAGCTACAGCCTCACGGCTCAAGCCAAATACGTTGGAAG 3120
Db 3460 AAAAAGCCCTTGATTTGACAGCTACAGCAGGAAAGGCTTAATGCCGAACATTTGGAAG 3519

Qy 3121 AATGTGACATTTAATGAGTGTGTTCAACTATCCACTGACAGACATCCCGTGT 3180
Db 3520 AATGTGACATTTGATGAGTGTGTTCAACTATCCACTGACAGACATCCCGTGT 3579
Qy 3181 CCAGGGGCTGAGCTCCGAGGAGTGAAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 3240
Db 3580 TCGAGGACTGAGCTCGAGGAGTGAAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 3639
Qy 3241 CTGTGGAAAGACACAGTTGTCTAGCTCTAGAGCGCTTCTATGACCCCTTGGCTGTT 3300
Db 3640 CTGTGGAAAGACACAGTTGTCTAGAGCGCTTCTATGACCCCTTGGCTGTT 3399
Qy 3301 AGTGTAAATTTGAGCAAGAGATAAAGACACGATGTGACAGTGGCTCGAGACAGCT 3360
Db 3700 AGTGTGCTGTGATGCAAAAGAAATTAAGCGATGATGTTGATGCTCGAGACAGCT 3759
Qy 3361 GGGCATCGTGTCTCAGAGCCCATCGTTTGTACGACAGATGTCGCGAGAACTTGCTTA 3420
Db 3760 GGGCATCGTGTCTCAGAGCCCATCGTTTGTACGACAGATGTCGCGAGAACTTGCTTA 3819
Qy 3421 TGGAGCAAAAGCCGGGTCTGATCAGTGAAGATTTGAGGACGCAAGAGGCCAA 3480
Db 3820 TGGAGCAAAAGCCGGGTCTGATCAGTGAAGATTTGAGGAGGAGGAGGAGGAGGAGGAG 3879
Qy 3481 CATACACACTTCACTGAGACACTCCCTGAGAAATACAAACACAGAGTAGAGACAAAG 3540
Db 3880 CATACACTGCTTCACTGAGACACTCCCTGAGAAATATATATACACTAAAGTAGAGACAAAG 3939
Qy 3541 AACCCAGCTCTGTGTTGGCCGAAGAACGCAATGTCATGAGCTCGCGCTTGTATAGCA 3600
Db 3940 AACCCAGCTCTGTGTTGGCCGAAGAACGCAATGTCATGAGCTCGCGCTTGTATAGCA 3999
Qy 3601 GCTCTATTTTGTGTTTGGATGAAGCTATACAGCTCTGATATACAGAAATGAAAAGGT 3660
Db 4000 GCTCTATTTTGTGTTTGGATGAAGCTATACAGCTCTGATATACAGAAATGAAAAGGT 4059
Qy 3661 TGTCCAAAGAGCCCTGCAAAAGCCAGAGAGGCGCACCTGATTTGATCGGCCACG 3720
Db 4060 TGTCCAAAGAGCCCTGCAAAAGCCAGAGAGGCGCACCTGATTTGATTTGATCGGCCACG 4119
Qy 3721 CTGTGACACATCTCGAATGAGATTAATAGTGGTTTGCAGATGGCAAGCAAGCA 3780
Db 4120 CTGTGACACATCTCGAATGAGATTAATAGTGGTTTGCAGATGGCAAGCAAGCA 4179
Qy 3781 GCATGGCACATTCACAGCTGCTGGCCAGAAAGCATTTTTCATGCTGACTGT 3840
Db 4180 GCATGGCACATTCACAGCTGCTGGCCAGAAAGCATTTTTCATGCTGACTGT 4239
Qy 3841 CCAGGCTGGAGCAAGCCCTAGTGAACCTGTGGCCATATGAGCTGTTAAATATTTTAAAT 3900
Db 4240 CCAGGCTGGAGCAAGCCCTAGTGAACCTGTAGATGATGATGTTAAATACTTTTAAAT 4299
Qy 3901 ATTGTTTAAACATGSCATTAATCAAGGTTAAAGAGTGGCACTACAGGAAAGCT 3960
Db 4300 ATT -TGTTAATATATGACATTTATTCAAAGTTAA -AGCAACACTTACAGAAAT 4352
Qy 3961 ATGTGAACCTACCTGTTTAACATTTCTGCTGACACTGAAAGATCTTCACCAAGTTGAG 4020
Db 4353 ATGTGAACCTACCTGTTTAACATTTCTGCTGACACTGAAAGATCTTCACCAAGTTGAG 4400
Qy 4021 AGTCCTCAAGATTTTAATTAAGAACCA -AAGAACATTTATCTGATGAGTAAAT 4078
Db 4401 -AGAGACTTCTATTTAAAGAACAGAGTGAAGACATCAAGTGGAGAGAAAT 4455
Qy 4079 ACTGTTTAAATGATTAATTAATTTATAGATTAATCAAGATGATTTTGTATTAAT 4138
Db 4456 CATAGTTTAACTGATTAATTAATTTATTAACAGATTAAGATGATTTTAAAGATTA 4515
Qy 4139 TGTATATATTTTGTATATTTT -ATTTGTAATCTAGCTGCTTGTGCAAGAT 4190
Db 4516 ATGTGTAATTTTGTATATTTTCCATTTTGAAGCTGATGACGCTGCTGCTAAAGAT 4575
```


QY	2346	TCACATTTTGGCAAAAGCTTGGGAGATGCTCCACTAAGCGGGCTTGATGCATGTTTTACAT	2
Db	3543	TCACATTTTGGCAAAAGCTTGAGATGCTCCACCAAGCGGCTCGATACATGTTTTCCAT	3
QY	2406	CCATCTCAGACAGATGTCAAGCTGGTTTGATGACCTTAAAAACACACTGAGCATTTGA	2
Db	3603	CCATGCTACAGACAGATGTAGTTGGTTTGATGACCTTAAAAACACACTGAGCATTTGA	3
QY	2466	CAACGAGGCTTGCCAAATGATGACGGCTCAATTTAAAGGGCTTTAGGTTCCAGGCTTGCTG	2
Db	3663	CTACGAGGCTTGCCAAATGATGATGTTGGTTTGATGACCTTAAAAACACACTGAGCATTTGA	3
QY	2526	TCATTACCCAGAAATATATGCAAACTTGGGACAGGCAATTAATTAATCTTAAATCTATGGTT	2
Db	3723	TAAATTACCCAGAAATATATGCAAACTTGGGACAGGAAATTAATTAATCTTAAATCTATGGTT	3
QY	2586	GGCAATTTAACATTTTACTTCTATAGCAATTTGTCCCATCATTTGCCATATACAGAGATTTGTTG	2
Db	3783	GGCAATTTAACATTTTACTTCTATAGCAATTTGTCCCATCATTTGCCATATATACAGAGATTTGTTG	3
QY	2666	AAATGAAATAGTTGTGTGGACACAGCACTGAAAAGATTAAGAAAAGACTTGAAGAGCACTGGGA	2
Db	3843	AAATGAAATAGTTGTGTGGACACAGCACTGAAAAGATTAAGAAAAGACTTGAAGAGCACTGGGA	3
QY	2706	AGATTCTCAAGACAGCAATCGAAAACCTTCCGAACTGTTGTTCTTTGACTCGGGAGACAGA	27
Db	3903	AGATGCTCTCAAGACCAATATAGAAAACCTTCCGAAACCGTTGTTCTTTGACTCAAGAGACAGA	39
QY	2766	AGTTTGATATCACTGTATGCACAGAGTTTGCATGATCATACAGAAAACCTCTTTAGGAAG	28
Db	3963	AGTTTGATATCAATGTATGCTCAGAGTTTGCAGGTATCATACAAAACCTCTTTAGGAAG	40
QY	2826	CACACATCTCGGGGCTCAATTTCTATCAACCAAGCAATGATATTTTTCATATGCTG	28
Db	4023	CACACATCTTTGGAAATTAATTTCTTCAACCAAGCAATGATATTTTTCATATGCTG	40
QY	2886	GCTGTTCCGTTTGGTGCCTACTTGGTGCCAATAGTTTCATGAACTTCCAGATGCTC	29
Db	4083	GATGTTCCGTTTGGAGCCTACTTGGTGCACATTAATCATATGACCTTTGAGGATGCTC	41
QY	2946	TTTTGATATCTCACTATTGTCTTTGGGCCATGGCAATGGGGCAAGCTCACTTCATTG	30
Db	4143	TGTTAGTATTTTCAAGCTGTTGTCTTTGGGCCATGGCCGTGGGGCAATGCTTCATTG	42
QY	3006	CTCTGACTATGCCAAAAGCAAGATATCAGACCCCAATCATATCATATGATCAATTAATAAAA	30
Db	4203	CTCTGACTATGCCAAAAGCAAAATATCAGACCCCAATCATATCATATGATCAATTAATAAAA	42
QY	3066	GCCTCTGATTTGACACTCAAGCCCTCAAGGCTCAAGCCAAATACGTTGGAAAGAAATG	312
Db	4263	CCCTTTGATTTGACACTCAAGCCCAAGGAGGAGGCTATATCCCAACACTTGGAAAGAAATG	432
QY	3126	TGACATTTAATGAGTGTGTTCAACTATCCCACTGCAGACAGACATCCCGGTGCTCAGG	318
Db	4323	TCACATTTGATGAAATTTGATTTCAACTATCCCAACCAAGGAGGAGGCTATATCCCAAGG	438
QY	3186	GAGTGAGCCTTCGAGGGAAGAAAGGGCCAGACAGCTGGCCCTGTAGTAGTACAGTGGCTGTG	324
Db	4383	GAGTGAGCCTTCGAGGGAAGAAAGGGCCAGACAGGCTGGCTGTGAGTAGTACAGTGGCTGTG	444
QY	3246	GGAAAGACACAGTTTTCAGTCTCCAGAGCCCTCTATGACCCCTTGGGCTGTGAGTGC	330
Db	4443	GGAAAGACACATGTTTCCAGTCTCCGAGGAGGCTTCTATGACCCCTTGGGAGGAAAGTGC	450
QY	3306	TAAATGATGGCAAGAGATTAAGCACTGAAATGTCCAGTGGCTCGAGACACACTGGGCA	336
Db	4503	TGCTTGATGGCAAGAAATTAAGGAGTGAATTTTGAAGTGTCCGAGACACACTGGGCA	456
QY	3366	TGCTGTCTCAGAGCCCATCCTGTTTGTACAGCAATTTGCCGGAACATTTGGCTTATGAG	342
Db	4563	TGCTGTCTCAGAGCCCATCCTGTTTGTACAGCAATTTCTAGAACATTTGCTTATTTGAG	462
QY	3426	ACAACAGCGGGGTGTATCACTAATGAAGATTTATGCAAGGACCCAGAGAGGCAACATATC	348

```

Db      4623  ||||||||||| || ||||| ||||| || ||||| ||||||||||| ||||| 4682
          ACACAGCGGGGTGTACAGAGAAAGATCGTAGGCGAGCAAGAGGCCACATAC
QY      3486  ACCACTTCATGAGACACCTCCCTGAGAAATACACACAGAGTAGAGACAAAGAACCC 3545
          | ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db      4683  ATGCCTCATGAGTACACTGCTAATTAATATAGCATAAAGTAGAGACAAAGAACTC 4742
          ||||||||||| ||||||||||| ||||||||||| |||||||||||
QY      3546  ACCTCTCGTGGGCCAAGAACAGCATATGCGATAGCTGCGGCTCTGTTAGACAGCTC 3605
          ||||||||||| ||||||||||| ||||||||||| |||||||||||
Db      4743  ACCTCTCGTGGGCCAAGAACAGCATATGCGATAGCTGCGGCTCTGTTAGACAGCTC 4802
          ||||||||||| ||||||||||| ||||||||||| |||||||||||
QY      3606  AATATTTCTTTTGGATGAAGCTACATCAGCTCTGATGATACAGAAAGTGAAGCTTGC 3665
          ||||||||||| ||||||||||| ||||||||||| |||||||||||
Db      4803  AATATTTCTTTTGGATGAAGCCACGCTGATGATACAGAAAGTGAAGCTTGC 4862
          ||||||||||| ||||||||||| ||||||||||| |||||||||||
QY      3666  AAGAACCCCTGACAAAGCCAGAGAGAGCCGACCTGCATTTGATGAGCCACCGCTGT 3725
          ||||||||||| ||||||||||| ||||||||||| |||||||||||
Db      4863  AAGAACCCCTGACAAAGCCAGAGAGAGCCGACCTGCATTTGATGAGCCACCGCTGT 4922
          ||||||||||| ||||||||||| ||||||||||| |||||||||||
QY      3726  CCACCATCCAGATGAGATTTAATAGTGTGTTTCAAGGCGCAAGTCAAGAGCATG 3785
          ||||||||||| ||||||||||| ||||||||||| |||||||||||
Db      4923  CCACCATCCAGATGAGATTTAATAGTGTGTTTCAAGATGCGAGAGCATG 4982
          ||||||||||| ||||||||||| ||||||||||| |||||||||||
QY      3786  GCACACATCAACAGCTCTGCGCCAGAAAGCATATATTTCCATGTCAGTCCAGG 3845
          |||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db      4983  GCACACATCAACAGCTCTGCGCCAGAAAGCATATATTTCCATGTCAGTCCAGG 5042
          |||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY      3846  CTGAGCAAAAGCGCTAGTGAAGTGTGCGCCATGAGCTGTTAAATATTTTAAAT 3900
          ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db      5043  CTGAGCAAAAGCGCGAGTGAAGTGTGAGTATGAGATGTTAAATCTTTTAAAT 5097
          ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||

RESULT 4
US-09-306-417-1
; Sequence 1, Application US/09306417
; Patent No. US20020103144A1
; GENERAL INFORMATION:
;   APPLICANT: Heinrich-Pette-Institut
;   TITLE OF INVENTION: Retroviral Gene Transfer Vectors
;   FILE REFERENCE: P50491
;   CURRENT APPLICATION NUMBER: US/09/306,417
;   EARLIER FILING DATE: 1999-05-06
;   EARLIER APPLICATION NUMBER: DE 198 22 115
;   EARLIER FILING DATE: 1998-05-08
;   NUMBER OF SEQ. ID NOS.: 17
;   SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 1
; LENGTH: 8630
; TYPE: DNA
; ORGANISM: Artificial Sequence
; FEATURE:
;   OTHER INFORMATION: Description of Artificial Sequence: proviral
;   FEATURE:
;   NAME/KEY: misc_feature
;   LOCATION: (1)..(160)
;   OTHER INFORMATION: plasmid backbone (pUC)
;   FEATURE:
;   NAME/KEY: misc_feature
;   LOCATION: (161)..(677)
;   OTHER INFORMATION: 5'-LTR
;   FEATURE:
;   NAME/KEY: 5'UTR
;   LOCATION: (532)..(1219)
;   FEATURE:
;   NAME/KEY:
;   LOCATION: (1220)..(5062)
;   OTHER INFORMATION: m4 mdr-1 cDNA
;   FEATURE:
;   NAME/KEY: misc_feature
;   LOCATION: (5215)..(5774)
;   OTHER INFORMATION: 3'-LTR
;   FEATURE:
```

```

; NAME/KEY: misc_feature
; LOCATION: (5775)..(8630)
; OTHER INFORMATION: plasmid backbone (pUC)
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (1)..(8630)
; OTHER INFORMATION: retroviral expression vector Sfbeta71m4
US-09-306-417-1
```

```

Query Match      75.3%; Score 3223.4; DB 10; Length 8630;
Best Local Similarity 89.4%; Pred. No. 0;
Matches 3483; Conservative 0; Mismatches 406; Indels 6; Gaps 1;
```

```

QY      6  GCGAGTGGGATGATGATCTGGAAGAGGCGCTAAGGGAGTGCAGAGAAACTTCTGGA 65
          ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db      1209  GCGGCGCGCGCATGATGATCTTGAAGGGGACCCGCAATGAGAGCAAGAAAGAACTTTT 1268
          ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY      66  AATGGGCAAAAAAAGTAAAAAATGAGAGAAAGAAAGAAAGAAAGAAAGAAAGAAAGT 125
          ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db      1269  TTAACCTGAAACATTAAGAGTGAAGAAAGTAAAGAAAGAAAGAAAGAAAGAAAGTAT 1328
          ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY      126  TTGCAATGTTTCCGCTATTCAAATTTGGCTTGATAGCTTGTATATTTGGTGGCAATAG 185
          ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db      1329  TTTCAATGTTTCCGCTATTCAAATTTGGCTTGACAAAGTTGTATATGTTGGGGAAGCTTTGG 1388
          ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY      186  CTGCATCATCATGAGAGCTGCACCTCCCTCATGATGCTGCTTTTGGAAACATGACAG 245
          ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db      1389  CTGCATCATCATGAGAGCTGCACCTCCCTCATGATGCTGCTTTTGGAAATGACAG 1448
          ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY      246  ATAGCTTGGCAATAGAGAGAAATTTCAAGAAACAAACCTTTCCAGTTATTAATTAAGAA 305
          ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db      1449  ATATCTTTGCAATGCAAGAAATTTA-----GAAAGTCTGATGCTCAACATCAGTATTA 1502
          ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY      306  GTATTAGCAACATTAACAACTTTTCATCAACCATCTGAGAGAGAGAAATGACACGATAG 365
          ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db      1503  GAAAGTATCATCATGATACAGAGGTTCTTCATGAAATCTGGAGAGACATGACCAATATAG 1562
          ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY      366  CCTATTATACAGTGGGATCGGTGCTGCGATGCTGATGCTGATGCTTCAATCAGAGTTTCAT 425
          ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db      1563  CCTATTATACAGTGGGAAATTTGCTGCGGAGTGTGCTGCTGCTTCAATCAGAGTTTCAT 1622
          ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY      426  TCTGTGCTCGGAGAGAGAGAGACATGATCAGAAATTAAGAAACAAATTTTCAATGCTA 485
          ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db      1623  TTTGGTGCCTGGAGAGCTGGAAGACAAATACACAAATTAAGAAACAGTTTTCATGCTA 1682
          ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY      486  TCATGCAACAGAGATGCTGCTGTTGACGTGACGTGACGTGGGAGCTTAACACCCGCG 545
          ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db      1683  TTAATGCAAGAGAGATGAGCTGTTGATGTGACGATGTTGGGAGCTTAACACCCGAG 1742
          ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY      546  TCACAGAGATGCTCCCAAAATCAATGAAGAAATGGCGCAAAATTTGGAATGTTCTTTC 605
          ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db      1743  TTACAGATGATGCTCTAAGATTAATGAATTAATGTTGTCACAAATTTGGAATGTTCTTTC 1802
          ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY      606  ACTCAATAGCAACATTTTTCACCGGTTTATAGGGGCTTACAGTGGTGAAGCTAA 665
          ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db      1803  AGTCAATGCAACATTTTTCACCTGGTTATAGATTAATACAGTGGTGAAGCTAA 1862
          ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY      666  CCCTTGTGATTTTGGCCATCAGCCCTGTTTGGCATTTTAAAGCCCATCTGGCAAGA 725
          ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db      1863  CCCTTGTGATTTTGGCCATCAGCTCTGTTTGGCATGTAAGCTGCTGCTGGGAAAGA 1922
          ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY      726  TACTATCTATTAATGATTAAGAACTTTGGCTTATGCAAAAGCTGAGAGTAGCTG 785
          ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db      1923  TACTATCTATTAATGATTAAGAACTTTAGCTTATGCTGAGAGTAGCTG 1982
          ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY      786  AAGAACTCTAGAGCAATCAAGAACTGTGATGCTTTGGAGAGCAAAAGAAAGAACTTG 845
          ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db      1983  AAGAGTCTTGGGCAATTAAGAACTGTGATGCTTTGGAGAGCAAAAGAAAGAACTTG 2042
          ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY      846  AAAGGTACAACAAAATTTAGAGAAAGCTAAAGAAATTTGGATTAAGAAAGCTATCAGG 905
          ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db      2043  AAAGGTACAACAAAATTTAGAGAAAGCTAAAGAAATTTGGATTAAGAAAGCTATCAG 2102
          ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
```

QY	906	CCAACTTTCTAATGTCGCGCTTTCTTATTTGATCTATGATCATATGCTGCTTCT	965
Db	2103	CCAAATATTTCTAATGTCGCTTCTTCTGCTGATCTATGATCTTATGCTCTGCTTCT	2162
QY	966	GGTATGGACCTCTGCTGCTCTCCAGTCAATATTTCTATTTGGACAAGTACTGCTCT	1025
Db	2163	GGTATGGACCTCTGCTGCTCTCTCAAGGGGAATTTCTATTTGGACAAGTACTGCTCT	2222
QY	1026	TCCTTTCTGATTAATTTGGGGCTTTTATGATTTGGACAGCATCCCAACATTTGAACAT	1085
Db	2223	TCCTTTCTGATTAATTTGGGGCTTTTATGATTTGGACAGCATCTCCAACTTTGAACAT	2282
QY	1086	TTGCAAAACGAGAGAGACAGCTTATGAATCTTCAAGTAAATTTGACATTAACCAAGCA	1145
Db	2283	TTGCAAAACGAGAGAGAGAGCTTATGAATCTTCAAGTAAATTTGACATTAACCAAGCA	2342
QY	1146	TTGACAGCTATTTGAGAGAGAGGACATTAACCAAGTAAATTTGAAGAAATTTGAATTC	1205
Db	2343	TTGACAGCTATTTGAGAGAGAGGACATTAACCAAGTAAATTTGAAGAAATTTGAATTC	2402
QY	1206	AAAATGTTCACTTCACTTACCTTCTCGAAAAGAGTTAAGATCTTAAAGGCTCAAC	1265
Db	2403	AAAATGTTCACTTCACTTACCTTCTCGAAAAGAGTTAAGATCTTAAAGGCTCAAC	2462
QY	1266	TGAAGGTCAGAGTGGGACAGACAGTGGGCTGTGGGACACAGTGGCTGGGGAAGACA	1325
Db	2463	TGAAGGTCAGAGTGGGACAGACAGTGGGCTGTGGGACACAGTGGCTGGGGAAGACA	2522
QY	1326	CGACCTGACAGTGTGACAGAGGCTCTATGACCCACAGATGGGATGGCTGATATGATG	1385
Db	2523	CGACCTGACAGTGTGACAGAGGCTCTATGACCCACAGATGGGATGGCTGATATGATG	2582
QY	1386	GACAGGACATTAAAGGCCATTAATGAAGCATCTTGGGAAATTAAGTGGTGGTGGTGC	1445
Db	2583	GACAGGACATTAAAGGCCATTAATGAAGCATCTTGGGAAATTAAGTGGTGGTGGTGC	2642
QY	1446	AGGAGCTGTGTTGTTGCCACACAGATAGCTGAAAACATTCGGTATGCGCGGAAATG	1505
Db	2643	AGGAGCTGTGTTGTTGCCACACAGATAGCTGAAAACATTCGGTATGCGCGGAAATG	2702
QY	1506	TCACCATGAGTGAATGAGAAAAGCTTTAAAGAGCCAAATGCTATGATTTATCATGA	1565
Db	2703	TCACCATGAGTGAATGAGAAAAGCTTTAAAGAGCCAAATGCTATGATTTATCATGA	2762
QY	1566	AACTACCTAATTAATTTGAACACTGCTGGTTGGAGAGAGAGGGGCCACAGCTGAGTGGAC	1625
Db	2763	AACTACCTAATTAATTTGAACACTGCTGGTTGGAGAGAGAGGGGCCACAGCTGAGTGGAC	2822
QY	1626	AGAAACAGAGATGCCCATTTGCTCGGGCCCTGTTGCAACCCCAAGATTTCTTGGCTGG	1685
Db	2823	AGAAACAGAGATGCCCATTTGCTCGGGCCCTGTTGCAACCCCAAGATTTCTTGGCTGG	2882
QY	1686	ATGAGCAACGTCAGCTGTCGACACTGAAAGTGAAGCAGTGGTTACAGTGGCCCTGATA	1745
Db	2883	ATGAGCAACGTCAGCTGTCGACACTGAAAGTGAAGCAGTGGTTACAGTGGCCCTGATA	2942
QY	1746	AGGCCGAAAAGGCGGACTACCATTTGATAGCTCATCGTTGTCTCACTGATGCTATG	1805
Db	2943	AGGCCGAAAAGGCGGACTACCATTTGATAGCTCATCGTTGTCTCACTGATGCTATG	3002
QY	1806	CCGATGTCATCTGCTGTTTTGATGAGTCAATTTGTGGAGAAAGAAATCATGATGAC	1865
Db	3003	CCGATGTCATCTGCTGTTTTGATGAGTCAATTTGTGGAGAAAGAAATCATGATGAC	3062
QY	1866	TCATGAAAAGAGAGGCAATTTACTTCAAACTGTACATGACACACAGCAGCAAGAAATGAA	1925
Db	3063	TCATGAAAAGAGAGGCAATTTACTTCAAACTGTACATGACACACAGCAGCAAGAAATGAA	3122
QY	1926	TTGAGTTGAAAAGGCGCTGCTGATTCCAAAGTGAAGTGAATGCTGCTC	1985
Db	3123	TTGAGTTGAAAAGGCGCTGCTGATTCCAAAGTGAAGTGAATGCTGCTGCTC	3182
QY	1986	CAAAAGATTTCAGGGTCCAGTTTAAATAAAAGAGATCAAGTCGAGAGATATACATGAC	2045

QY	3183	CAAAATGATTCAGATTCAGATTCATTAATTAAGAAAAGTCAAGTCGAGAGTGTCCGTGGAT	3242
QY	2046	CACAAGGCCAAGACAGAAAGCTTGTGTACAAAAGAGACTTGAATGAGAAATGTACTCCAG	2105
Db	3243	CACAAGGCCAAGACAGAAAGCTTGTGTACAAAAGAGACTTGTGAATGAGAAATGTACTCCAG	3302
QY	2106	TTTCTCTTGGAGGATTTGAAAGCTGAAGCTCACTGAATGGCCCTTATTTGGTGGTGA	2165
Db	3303	TTTCTCTTGGAGGATTTGAAAGCTGAAGCTCACTGAATGGCCCTTATTTGGTGGTGA	3362
QY	2166	TATTTTGTCTATTAATAACAGAGAGGCTCGCAACAGATTTTCAATTAATTTTCAAGGA	2225
Db	3363	TATTTTGTCTATTAATAACAGAGAGGCTCGCAACAGATTTTCAATTAATTTTCAAGGA	3422
QY	2226	TTATAGGATCTTTACCCGAGATGAGATCCTGAAACAAAACAGACAAATATGATACATG	2285
Db	3423	TTATAGGATCTTTACCCGAGATGAGATCCTGAAACAAAACAGACAAATATGATACATG	3482
QY	2286	TTTCTGATTTGTTCTGATGCTTGGAAATTTATTTCTTATTAATTTTCTCCAGGCT	2345
Db	3483	TTTCTGATTTGTTCTGATGCTTGGAAATTTATTTCTTATTAATTTTCTCCAGGCT	3542
QY	2346	TCACATTTGGCAAAAGCTGGGAGATCCTCACTAAGCGGCTGATACATGTTTCAGAT	2405
Db	3543	TCACATTTGGCAAAAGCTGGGAGATCCTCACTAAGCGGCTGATACATGTTTCAGAT	3602
QY	2406	CCATGCTGACAGAGATGTCAGCTGTTGATGATGACCTTAAAAACACACTGAGCATTTGA	2465
Db	3603	CCATGCTGACAGAGATGTCAGCTGTTGATGATGACCTTAAAAACACACTGAGCATTTGA	3662
QY	2466	CAACACAGGCTTGCCAAATGATGCGGCTCAAGTTAAAGGGGCTTAAAGTTCAGAGCTGGCG	2525
Db	3663	CAACACAGGCTTGCCAAATGATGCGGCTCAAGTTAAAGGGGCTTAAAGTTCAGAGCTGGCG	3722
QY	2526	TCATTTCCGAGATTTAAGCAATCTTGGACAGCATTTATTTATCTTAATCTAATGCTT	2585
Db	3723	TCATTTCCGAGATTTAAGCAATCTTGGACAGCATTTATTTATCTTAATCTAATGCTT	3782
QY	2586	GGCAATTAACACTTTTACTCTTGAACATTTGACCATATTTGACATGACAGAGTGGTG	2645
Db	3783	GGCAATTAACACTTTTACTCTTGAACATTTGACCATATTTGACATGACAGAGTGGTG	3842
QY	2646	AAATGAAATGTTGCTGAGACAGACACTGAAGATTAAGAAAGACTGAAGAGCTGGGA	2705
Db	3843	AAATGAAATGTTGCTGAGACAGACACTGAAGATTAAGAAAGACTGAAGAGCTGGGA	3902
QY	2706	AGATTTCTACAGAGCCATCGAAAGCTTCCGAACCTGTTGTTCTTACTCTGAGACAGA	2765
Db	3903	AGATTTCTACAGAGCCATCGAAAGCTTCCGAACCTGTTGTTCTTACTCTGAGACAGA	3962
QY	2766	AGTTGAAATATATGATATGACAGAGTGTGCAAGTACCATACAGAAACTTGTGAGAAAG	2825
Db	3963	AGTTGAAATATATGATATGAC	

Db 4263 CCCCTTTGATGACAGTACAGCAGGAAAGCCCTAATGCCAGACATTTGGAAGAAATG 4322
Qy 3126 TGACATTTAATAGAGTCGCTGTTCACTATCCACTCGACACAGACATCCCGTGTCCAGG 3185
Db 4323 TCACATTTGTAAGTGTGATTTCACTATCCACACCCGACCGACATCCCGTGTCCAGG 4382
Qy 3186 GGCTGAGCCTCGAGTGAAGAGGCGCAGAGCTTGCCCTGTAGTATGAGTGGCTGTG 3245
Db 4383 GACTGAGCCTCGAGTGAAGAGGCGCAGAGCTTGCCCTGTAGTATGAGTGGCTGTG 4442
Qy 3246 GGAAGAGCAGAGTGTGCTGCTGATGAGAGCCTTATGACCCCTTGGCTGTGATGTC 3305
Db 4443 GGAAGAGCAGAGTGTGCTGCTGATGAGAGCCTTATGACCCCTTGGCTGTGATGTC 4502
Qy 3306 TAATTTGATGGAAGAGATTAAGACCTGATATGTCAGAGCTGCTCGACACACCTGGGCA 3365
Db 4503 TGGTGTATGGAAGAAATTAAGCAGTGAATGTCAGAGCTGCTCGAGACACCTGGGCA 4562
Qy 3366 TCGTGTCTCAGAGAGCCCATCTGTTGACTGACGATTTGCTGAGAACATTTGCCATGAG 3425
Db 4563 TCGTGTCTCAGAGAGCCCATCTGTTGACTGACGATTTGCTGAGAACATTTGCCATGAG 4622
Qy 3426 ACAACAGCCGGGTGTATCATATGAAGATTAATGCAAGCAGCAGCAGGAGGCGCAACATAC 3485
Db 4623 ACAACAGCCGGGTGTGTACAGAGAGATGCTGAGGAGCAGAAAGGAGGCGCAACATAC 4682
Qy 3486 ACCACTTCATGAGACACTCCCTGAGAAATTAACAACACAGATGAGAGACAAAGAACCC 3545
Db 4683 ATGCCCTCATGAGTCACTGCTATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 4742
Qy 3546 AGCTCTGTGTGCGCAGAAACAGGCGCATTTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 3605
Db 4743 AGCTCTGTGTGCGCAGAAACAGGCGCATTTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 4802
Qy 3606 ATATTTTCTTTTGTATTAAGTACATCACTGCTGATGATGATGATGATGATGATGATGAT 3665
Db 4803 ATATTTTCTTTTGTATTAAGTACATCACTGCTGATGATGATGATGATGATGATGATGAT 4862
Qy 3666 AAGAAGCCCTGAGCAAAAGCCAGAGAGGCGCAGCTGCTGATGATGATGATGATGATGAT 3725
Db 4863 AAGAAGCCCTGAGCAAAAGCCAGAGAGGCGCAGCTGCTGATGATGATGATGATGATGAT 4922
Qy 3726 CCACCATCCAGATGAGATTTAATAGTGTGTTTCAAGATGGAAGTCAAGAGCATG 3785
Db 4923 CCACCATCCAGATGAGATTTAATAGTGTGTTTCAAGATGGAAGTCAAGAGCATG 4982
Qy 3786 GCACATATCAAGAGTGTGCGCCAGAAAGGATCTATTTTTCATGCTCAGTGTCCAGG 3845
Db 4983 GCACATATCAAGAGTGTGCGCCAGAAAGGATCTATTTTTCATGCTCAGTGTCCAGG 5042
Qy 3846 CTGAGCAAAAGCCTATGATGATGCTGATGATGATGATGATGATGATGATGATGATGAT 3900
Db 5043 CTGAGCAAAAGCCTATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 5097

RESULT 5
US-09-866-866A-1
; Sequence 1, Application US/09866866A
; Patent No. US20020102244A1
; GENERAL INFORMATION:
; APPLICANT: Sorrentino, Brian
; APPLICANT: Schubert, John
; TITLE OF INVENTION: A Method of Identifying and/or Isolating Stem Cells
; FILE REFERENCE: 1340-1-021CIP2
; CURRENT APPLICATION NUMBER: US/09/866,866A
; PRIOR FILING DATE: 2001-08-30
; PRIOR APPLICATION NUMBER: 09/584,586
; PRIOR FILING DATE: 2000-05-31
; PRIOR APPLICATION NUMBER: PCT/US99/11825
; PRIOR FILING DATE: 1999-05-27
; PRIOR APPLICATION NUMBER: 60/086,988
; PRIOR FILING DATE: 1998-05-28
; NUMBER OF SEQ ID NOS: 27

; SOFTWARE: PatentIn version 3.0
; SEQ ID NO 1
; LENGTH: 3860
; TYPE: DNA
; ORGANISM: Homo sapiens
US-09-866-866A-1
Query Match 74.8%; Score 3202; DB 10; Length 3860;
Best Local Similarity 89.6%; Pred. No. 0;
Matches 3455; Conservative 0; Mismatches 395; Indels 6; Gaps 1;
Qy 17 ATGATCTCTGAGGAGGCGCGTAAGGGAGTGCAGAGAACTTCTGGAATATGGCAAA 76
Db 1 ATGATCTCTGAGGAGGCGCGTAAGGGAGTGCAGAGAACTTCTGGAATATGGCAAA 60
Qy 77 AAAATGAAAAAATGAGAAAGAAAGAAAGAAAGAAAGAAAGAAAGAAAGAAAGAAAG 136
Db 61 AAAATGAAAAAATGAGAAAGAAAGAAAGAAAGAAAGAAAGAAAGAAAGAAAGAAAG 120
Qy 137 CGCTATTTCAATTTGCTGATAGTGTATATATGTTGTTGGGAGCAATGGCTGCCATATC 196
Db 121 CGCTATTTCAATTTGCTGATAGTGTATATATGTTGTTGGGAGCAATGGCTGCCATATC 180
Qy 197 CATGAGCTGACCTCCCTCTCATGATGCTGCTTTTGAAGACATGACAGATAGTTTGA 256
Db 181 CATGAGCTGACCTCCCTCTCATGATGCTGCTTTTGAAGACATGACAGATAGTTTGA 240
Qy 257 AATGAGGAATTTCAAGAAACAAACCTTTTCCAGTTATTAATTAAGAAATTTACGAA 316
Db 241 AATGAGGAATTTA-----GAAAGATCTGATGTCAAACTACATTAATGAAGTAAATC 294
Qy 317 AATGAGGAATTTTCAATCAACATCTGAGAGAGAAATGACCGATTAATTAATTAAT 376
Db 295 AATGATACAGGCTCTTCAATGAATGAGAGAGAACTGACGACGATATGCTATTAAT 354
Qy 377 AGTGGATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 436
Db 355 AGTGGATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 414
Qy 437 GCAGCAGGAAGAGATCTCAAAATTTGAAGAAATTTTTCATGCTATCATGACGAG 496
Db 415 GCAGCAGGAAGAGATCTCAAAATTTGAAGAAATTTTTCATGCTATTAATTAATTAAT 474
Qy 497 GAGATTTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 556
Db 475 GAGATTTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 534
Qy 557 GTCTCCAAATCAATGAAGAAATTTGCGCAAAATTTGAAGAAATTTTTCATGCTATCA 616
Db 535 GTCTCCAAATCAATGAAGAAATTTGCGCAAAATTTGAAGAAATTTTTCATGCTATCA 594
Qy 617 ACATTTTTCACCGCTTTATATGATGAGGTTTACAGTGTGGAAGCTTAACCTTTGAT 676
Db 595 ACATTTTTCACCGCTTTATATGATGAGGTTTACAGTGTGGAAGCTTTTACCTTTGAT 654
Qy 677 TTGGCATGAGCCTGCTTTTCTGAGCTTTCAGCCGCTGCTGCTGCTGCTGCTGCTGCTGCT 736
Db 655 TTGGCATGAGCCTGCTTTTCTGAGCTTTCAGCCGCTGCTGCTGCTGCTGCTGCTGCTGCT 714
Qy 737 TTACTGATTAAGAACTCTTGGCTTGAAGCAAAAGCTGAGCAAGTATGATGATGATGAT 796
Db 715 TTACTGATTAAGAACTCTTGGCTTGAAGCAAAAGCTGAGCAAGTATGATGATGATGAT 774
Qy 797 GCAGCATGAGCACTGATGATGCTTTGAGAGCAAAAGAAAGAAAGAAAGAAAGAAAG 856
Db 775 GCAGCATGAGCACTGATGATGCTTTGAGAGCAAAAGAAAGAAAGAAAGAAAGAAAG 834
Qy 857 AAAATTTTGAAGAGTAAAGAAATTTGGATTAAGAAAGATTAACAGGCAATTTCT 916
Db 835 AAAATTTTGAAGAGTAAAGAAATTTGGATTAAGAAAGATTAACAGGCAATTTCT 894
Qy 917 ATGAGCCTGCTTCTATGATCTATGATCAATCATGCTGCTGCTTCTGCTGATGGAGCC 976
Db 894 ATGAGCCTGCTTCTATGATCTATGATCAATCATGCTGCTGCTTCTGCTGATGGAGCC 976

Query Match	Best Local Similarity	Matches 3454; Conservative	74.8%; Score 3200.4; Pred. No. 0; Mismatches 396; Indels 6; Gaps	DB 10; Length 3860;
17	ATGGATCCTTAAGAGAGCCCTTAAGGGGAGTGCAGAGAAAGCACTTCTGGAATAATGGCGAA	76	11	
1	ATGGATCTTAAGAGGCGACGCAATGGAGAGCGAAAGAAAGAACTTTTAAACAGTAC	60	11	
77	AAAAGTAAAAAATGAGAGAAAGAAAGAAACCAGCTGTACAGCGTTTGCAATGTT	136	11	
61	AATAAAGTGAATAAGTAAAGAAAGAAAGAAACCACTGTAGGTATTTTCAATGTT	120	11	
137	CGCTATTCAATGGCTTGTATAGGTGTATATGTTGGGGGACAAATGGCGCATATC	196	11	
121	CGCTTTCAATTTGGCTTGGACAGTGTATATGTTGGTGGAACTTTGGCGCCATATC	180	11	
197	CATGAGGTGCACCTCTCATATGATGCTGTTTTGGAACATGACAGTATGCTTGA	256	11	
181	CATGGGCTGCACCTTCCCTCATATGCTGTGTTGGAGAAATGACAGATATCTTGA	240	11	
257	AATGACGAAATTTCAAGAAACAAACCTTTCCAGTTATATTAATGAAGATTTACGAC	316	11	
241	AATGACGAAATTTA-----GAAATCTATGTCCAAACATCACTATATGAGTATATC	294	11	
317	AATACACAATTTCTATCAACCATCTGGAGAGAAATGACACCGTATATATATAC	376	11	
295	AATGATACAGGGTTCTCATGAATCTGGAGAAACATGACACGATATATCTTATATAC	354	11	
377	AGTGGATGCGTGGCGGGGCTGGTGGTGGCTTACATCCAGGTTTCATTTGGTCTG	436	11	
355	AGTGGATTTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGG	414	11	
437	GCAGCAGAGAAACAGATACTCAAAATTTGAAACAAATTTTATATGTCATGACACAG	496	11	
415	GCAGCTGAGAAACAAATACCAAAATTTGAAACAAATTTTATGCTATATGCTATATAC	474	11	
497	GAGATTTGGGCTGTTTACGTGATACGTTGGGAGCTTAAACCCGCTTACAGATGAT	556	11	
475	GAGATTTGGGCTGTTTATGATGATGATGATGATGATGATGATGATGATGATGATG	534	11	
557	GCTCTCAAAATCAATAGAGATTTGGGCAAAATTTGAATGTTCTTCACTCAATAGCA	616	11	
535	GCTCTCAAAATCAATAGAGATTTGGGCAAAATTTGAATGTTCTTCACTCAATAGCA	594	11	
617	ACATTTTTCACCGTTTTATAGTGGGTTTACAGTGGTGGTGGTGGTGGTGGTGGTGG	676	11	
595	ACATTTTTCACCGTTTTATAGTGGGTTTACAGTGGTGGTGGTGGTGGTGGTGGTGG	654	11	
677	TTGGCATCAGCCCTGTTTGGACCTTACGCGCATCTGGGCAAAAGATCTATCTTCA	736	11	
655	TTGGCATCAGCCCTGTTTGGACCTTACGCGCATCTGGGCAAAAGATCTATCTTCA	714	11	
737	TTTACTGATTAAGAACTCTTGGCTTATGCAAAAGCTGGAGCAAGTATGCTTGA	796	11	
715	TTTACTGATTAAGAACTCTTGGCTTATGCAAAAGCTGGAGCAAGTATGCTTGA	774	11	
797	GCAGCAATCAACGATGATGCTTGGAGGAGCAAAAGAAAGAACTTGAAGGTACAC	856	11	
775	GCAGCAATCAACGATGATGCTTGGAGGAGCAAAAGAAAGAACTTGAAGGTACAC	834	11	
857	AAAAATTTAGAGAACTTAAGGAATTTGGGATTAAGAAAGCTATCAGGCGCAACTTCT	916	11	
835	AAAAATTTAGAGAACTTAAGGAATTTGGGATTAAGGAAGCTATTAACAGCAATTTCT	894	11	
917	ATATGTCGCGCTCTTATATATCATGATCATATGCTTGGCTTTTGTGATGGAC	976	11	
895	ATATGTCGCTCTTCTCTCTATATCATCTTATCTCTGCGCTTGTGATGGAC	954	11	

Db	2035	GACAGAAAGGCTTAGTACCAAGAGGCGTGGATGAAAGTATACCTCAGTTCTCTTTGG	2096
QY	2117	AGGATTTCGAAGCTGAACCTCAACGATGAGCCTTATTTTGTGGTGTATATTTGTGCT	2176
Db	2095	AGGATTATGAAGCTAAATTTTAACGATGAGCCTTATTTTGTGGTGTATATTTGTGCTC	2154
QY	2177	ATTATTAACGAGGCTCTGCACACGATTTTTCATTAATATTTTTCACAGATTTATAGGATC	2236
Db	2155	ATTATTAATGAGGCGCTGCACACGATTTTTCATTAATATTTTTCACAGATTTATAGGATC	2214
QY	2237	TTTACCCGAGATGAGATCCTGAAACAAACGAGAAATAGTAACATGTTTGTGTATG	2296
Db	2215	TTTACGAAATTTATATCTCTGAAACAAACGAGAAATAGTAACATGTTTGTGTATG	2274
QY	2297	TTTCTACTCTTGGAAATATTTCTTTTATTACATTTTTCCTCCAGGCTTCACATTTGGC	2356
Db	2275	TTTCTACCCCTTGGAAATATTTCTTTTATTACATTTTTCCTCCAGGCTTCACATTTGGC	2334
QY	2357	AAAGCTGGGAGATCCTCCTCAAGGGCGTTCATTCATGTTTTCAGATCCATGCTAGA	2416
Db	2335	AAAGCTGGGAGATCCTCCTCAAGGGCGTTCATTCATGTTTTCAGATCCATGCTAGA	2394
QY	2417	CAGAGTGCAGCTGTGTTTGATGACCCCTAAAAACACATGAGATGAGATGACACAGGCTT	2476
Db	2395	CAGAGTGCAGCTGTGTTTGATGATACCCCTAAAAACACCTGAGATGAGATGACACAGGCTT	2454
QY	2477	GCCAAATGATGGGCGTCAAGTTTAAGGGCGTATAGGTTCCAGGCTTGCTCATTTACCAG	2536
Db	2455	GCCAAATGATGGGCGTCAAGTTTAAGGGCGTATAGGTTCCAGGCTTGCTCATTTACCAG	2514
QY	2537	AATATAGCAAAATCTTGGGACAGGCAATATTTATTCCTATTCATGTTGGCAATTACA	2596
Db	2515	AATATAGCAAAATCTTGGGACAGGCAATATTTATATATCTTCACTATGTTGGCAAACTACA	2574
QY	2597	CTTTTACTCTTAGCAATTTGTACCATATTTGCATATAGCAGAGATTTGTAATGAAATG	2656
Db	2575	CTTTTACTCTTAGCAATTTGTACCATATTTGCATATAGCAGAGATTTGTAATGAAATG	2634
QY	2657	TTTGTCTGACAAAGCACTGAAAGATTAAGAAAGACCTTGAAGGAGCTGGGAAATGCTTCA	2716
Db	2635	TTTGTCTGACAAAGCACTGAAAGATTAAGAAAGACCTTGAAGGAGCTGGGAAATGCTTCA	2694
QY	2717	GAGCAATTCGAAACTTCCGAATGTTTGTCTTTTGACTCGGAGCAGAAATTTGAATAC	2776
Db	2695	GAGCAATTCGAAACTTCCGAATGTTTGTCTTTTGACTCGGAGCAGAAATTTGAATAC	2754
QY	2777	ATGTATGCACAGATTTGCAAGTACCATTCAGAAACTCTTTGAGAAAGCAACATCTTC	2836
Db	2755	ATGTATGCACAGATTTGCAAGTACCATTCAGAAACTCTTTGAGAAAGCAACATCTTC	2814
QY	2837	GGGGTCTCATTTTCTATCCACCAGCAATGATATTTTTCCTATGCTGCGTGTTCGG	2896
Db	2815	GGAAATTCATTTTCTTCCATCCACAGCAATGATATTTTTCCTATGCTGCGATGTTTCGG	2874
QY	2897	TTTGTGCTACTTGGTGTGCAAAATAGTTCATGAACTTTCAGGATGTTCTTGTGTATTC	2956
Db	2875	TTTGTGCTACTTGGTGTGCAAAATAGTTCATGAACTTTCAGGATGTTCTTGTGTATTC	2934
QY	2957	TCAGCTATTGTTTGTGTGCCATGGCAGTGGGAGGTGAGCTCATTTTGTGCCGCAAT	3016
Db	2935	TCAGCTATTGTTTGTGTGCCATGGCAGTGGGAGGTGAGCTCATTTTGTGCCGCAAT	2994
QY	3017	GCCAAAGCCAAAGTATTCAGACGCCAGCTCATATGATTCATTAAGAAAGCCCTCGATT	3076
Db	2995	GCCAAAGCCAAAGTATTCAGACGCCAGCTCATATGATTCATTAAGAAAGCCCTCGATT	3054
QY	3077	GACACTACAGCCTCAGCGGCTCAAGCCAAATACGTTGGAAGAAATGACATTTAAT	3136
Db	3055	GACACTACAGCAGGAAAGGCTTAATCGCAACATTTGGAAGAAATGTCACATTTGGT	3114
QY	3137	GAGGTCGTTCACTATCCACTGACACAGACATCCCGTGTCTCAGGCGTAGGCTTC	3196

Db	943	ICCTGGTCACTTCGCAAGAAATACCTGATTTGGACGAAGTGCTCACTGTCTCTTTCCGTG	1002
QY	1037	TTAATTGGGGCTTTTATGATTTGAGCAGGATCCCAACATTTGAAGCATTTTGCAAACGCA	1096
Db	1003	TTAATTGGAGCATCATGCTGTGTGACAGGCATCTCCAAATATTGAAGCCTTCGCCAAATGCA	1062
QY	1097	AGAGGACGACGTTATGAAATCTTCAGATTAATTGCAATTAACCAAGCATGACAGCTAT	1156
Db	1063	CGAGGAGCAGCTTATGAAGTCTTCAAAATTAATTGATATTAAGCCCATGATGACAGCTTC	1122
QY	1157	TCGAAGAGTGCACATTAACACGATTAATTAAAGGAATTTTGGATTTCAAAAATGTTTAC	1216
Db	1123	TCAAAGAGTGGGCACAAACGACACACATACAGGAATCTGGAAATTTAAGAAATATTTCAC	1182
QY	1217	TTTCAGTTACCTTCTCTGAAAAGAAAGTTAAGATCTTAAAGGGTCTCAACTGGAAGTTCAG	1276
Db	1183	TTTCAGTTACCCATCTCGAAAAAGAAAGTTTCAGATCTTGAAGGGCCTCAATCTGAAGGTGAAG	1242
QY	1277	AGTGGGACGACAGTGGGCGCTGGTTGGGACAGTGGCTGCGGGAAAGACACAGCCGTGACG	1336
Db	1243	ACGGGACAGACGGTGGCCCTGTGGGCAACGTTGGCTGTGAAAAAGCAACATGTCTCAG	1302
QY	1337	CTGATGCGAGAGCTCTATGACCCCAACAGATGGCATGTCTGTATTATGAGCAGGACATT	1396
Db	1303	CTGAATGCAAAAGCCTCTACGACCCCCCTAGATGGCATGTGCTAGTTCACAGCAGAGCATTC	1362
QY	1397	AGGACCATTAATTAAGGACATCTTCGGGAAATTAATCTGTGTGTGTGATGATCAGAGCCCTGTG	1456
Db	1363	AGAACCATCATGATGAGGTATCTGAGGGAATATATTGTGTGTGTGATGAGAACCTGTGG	1422
QY	1457	TTGTGTTGCCACACGATAGCTGAAAAACATTCGCTATGGCGCGGAAATATGCATGAT	1516
Db	1423	CTGTGTTGCCACACGATAGCTGCGGAAACATTCGCTATGGCGCGGAAATATGCATGAT	1482
QY	1517	GAGATTGAGAAGCTGTTAAGGAAGCAATGCTATGATTTATCATGAACTACTAAT	1576
Db	1483	GAGATTGAGAAGCTGTTAAGGAAGCAATGCTATGATTTATCATGAACTACTAAT	1542
QY	1577	AAATTTACACTGTGTTGGAGAGAGAGGGGCCACAGCTAGTGAAGTGTGAGACAGAAACAGGA	1636
Db	1543	CAATTTGACACCTGTGTGTGTGAGAGAGGGGCGACGTGATGTGGGGGACACAGAAACAGGA	1602
QY	1637	ATCGCCATTCTGTGGGCGCCGTGGTGGCCAAACCCCAAGATTTCTTCTGTGATGAGCAACG	1696
Db	1603	ATCGCCATTCTGTGGGCGCCGTGGTGGCCAAATCCCAAGATCTTCTTGTGTGAGAGAGCCACC	1662
QY	1697	TCAGCTCTGACACTGTAAGAGTGAAGCAGTGTCTCAGTGTGGCCCTGGAATGAAGCCACGAAA	1756
Db	1663	TCAGCCCTCGAATACAGAAAGTGAAGCTGTGGTTACGGCCGACCTGGATMAAGCTAGAGAA	1722
QY	1757	GGCGGACATACCAATTTGTATACCTCATCTGTTGTCTTACAGTTTGGTAATGGCATGTCAAT	1816
Db	1723	GGCGGACACCACTTTGTATGTATGTACTCATGTGTGTCTCTACCGTTCTGTAATGCTGACGTCA	1782
QY	1817	GCTGGTTTGTATGATGAGAGTCAATGTGTGAGAGAAAGAAATATGATGTAACATCAATGAAGAG	1876
Db	1783	GCTGGTTTGTATGATGAGTGTGTCTCATTTGTGGACGACAGGAATCATGATGAGCTCATGAGAGAA	1842
QY	1877	AAGGCGATTTACTTAAACTGTACATGTGACACAAAGAGGAATGAATTTGATTAGAA	1936
Db	1843	AAGGCGATTTACTTAAACTGTGTATGACACGACACAGAGAAATGAATTTGATTAGGA	1902
QY	1937	AATGCCACTGTGATCCAAAGTGAAGAGTGAATCCCTTGGAAATGTGTCCAAAGATTGA	1996
Db	1903	AATGAAGCTTGTAAATCTTAAGGATGAATTTGATTAATTTAGACATGTCTTCAAAAGTTGA	1962
QY	1997	GGGTGCACATTTAATAAAAAAGAAATCACTGCGAGAGATATACATGACACACAGGCCAA	2056
Db	1963	GGATCCACATCTAATTAAGAAAGAAATCACTGCGAAAGCATCTGTGACACATGACCAA	2022
QY	2057	GACAGAAAGCTTGTAGAAAAGAGACTTGAATGAAGAAATGTACCTCAAGTTCTCTTCTGG	2116
Db	2023	GACAGGAAGCTTGTATACCAAGAGGCCCTGTGATGAAGATGTATCTTCACACTCTTCTTGG	2082

QY	2117	AGGATTTGTGAAGCTGAACCTCAACTGATGAGCCTTATTTTGTGGTGGTATTTATTTGGCT	217
Db	2083	CGGATTCCTGGAAGTGTGAATTCACCTGATGAGCTTATTTTGTGGTGGTATTTCTGGC	214
QY	2177	ATTATTAAGGAGGCGCGCAACAGCATTTTCCATATATTTTTCACAGATTTATAGGATC	223
Db	2143	ATATATTAATGAGGCTTACAGCCAGCATTCCTCGTAAATATTTTTCAAAAGTTGAGGGTT	220
QY	2237	TTTTACCGGAGATGAGGATCCTGTGAACAAAACGACATAGTAAACAATGTTTCTGTATG	229
Db	2203	TTTACAAATGTGGTCCCGCCCTGAACCCAGCGGAGAACAGCAACTGTGTTTCTGTGTTG	226
QY	2297	TTTTCTAGTCCCTTGGATTTATTTCTTTTATTTACATTTTTCCTCAGGGCTTCACATTTGGC	235
Db	2263	TTTTCTATTCCTTGGGATCATTTCTTTTCATTTTACATTTTTCCTCAGGGCTTCACATTTGGC	232
QY	2357	AAAGCTGGGGAGATCCTCACTAAGGGGCTTCATACATGATGGTTTTCAGATCCATGCTGGA	241
Db	2323	AAAGCTGGAGAGATCCTCAACCAAGGACTCCGATACATAGGTTTTCATAATTCATGCTGGA	238
QY	2417	CAGGATGTAGCTGGTTGTGATGACCCCTAAAAACACCACTGGAGCATTTGACAAACAGGCTT	247
Db	2383	CAGGATGTAGCTGGTTGTGATGACCCCTAAAAACACCACTGGAGCATTTGACAAACAGGCTT	244
QY	2477	GCCATATGATGGCGCTCAAGTTAAAGGGGCTTTAGGTTCCAGGCTTGGCTGATTAACCGAG	256
Db	2443	GCCAAAGATGGCTGCTCAAGTGAAGGGGCTTACAGGGCTTCAGGCTTGGCTGATTAATTCAG	250
QY	2537	AATATAGCAATCTTGGGACAGGCATTTATTTATCTTAACTGATGATGGTTGGCATTTTACA	259
Db	2503	AACATAGCAAACTTGGGACAGGCATTTATTTATCTTAACTGATGATGGTTGGCATTTTACA	256
QY	2597	CTTTTACTCTTAGCAATTTGTACCATCATTTGACATTTACAGAGATTTGTTGAATGAATAG	265
Db	2563	CTTTTACTCTTAGCAATTTGTACCATCATTTGACATTTGACATTTGTTGAATGAATAG	262
QY	2657	TTTGTCTGGACAGACACTGAAAAGATTAAGAAAGAGCTTAAGAGAGCTGGAGATTTGCTACA	271
Db	2623	TTTGTCTGGACAGACACTGAAAAGATTAAGAAAGAGCTTAAGAGAGCTGGAGATTTGCTACA	268
QY	2717	GAACCCATTCGAAAACCTTCGGAACGTGTGTTCTTTGACTGGGAGACAGATTTGAATAC	276
Db	2683	GAACCATTTGAAAACCTTCGGAACGTGTGTTCTTTGACTGGGAGACAGATTTGAATAC	274
QY	2777	ATGTATGACAGAGATTTGCAAGTATACATACAGAAACCTTTTGGAGAAACACACATCTTC	283
Db	2743	ATGTATGACAGAGATTTGCAAGTATACATACAGAAACCTTTTGGAGAAACACACATCTTC	280
QY	2837	GGGGTTCATTTTCTATTCACACCCAGGCAATGATGATTTTTCCTATGCTGGCTTTCCGG	289
Db	2803	GGGATCACGTTCTCTCTTCACACCCAGGCAATGATGATTTTTCCTATGCTGGCTTTCCGG	286
QY	2897	TTTTGTGCCCTACTTGTGTGCAATGATGATGATGATGATGATGATGATGATGATGATGATG	295
Db	2863	TTTTGTGCCCTACTTGTGTGCAATGATGATGATGATGATGATGATGATGATGATGATGATG	292
QY	2957	TCGATATTTGCTTTTGTGTGCTCATGGCAGTGGGGCAGGCTCAGTTCTTTGCTCCTGACTAT	301
Db	2923	TCGATATTTGCTTTTGTGTGCTCATGGCAGTGGGGCAGGCTCAGTTCTTTGCTCCTGACTAT	298
QY	3017	GCCAAAGCCAAAGTATCAGACAGCCCAAGCATCATGATCATTTGAAAAAGCCCTCTGATT	307
Db	2983	GCCAAAGCCAAAGTATCAGACAGCCCAAGCATCATGATCATTTGAAAAAGCCCTCTGATT	304
QY	3077	GACACGTACAGCCCTCAGCGGCTCAAGCCAAATACGTTGAGAGGAAATGTGACATTTTAAT	313
Db	3043	GACACGTACAGAGCCCAAGGCTTAAAGCCGAATATGTGGAAGAAATGTGACATTTTAAT	310
QY	3137	GAGGTCGTGTTCAATTTATTCACATGACACACATCCCGCTGCTCCAGGGCTGAGGCTC	319
Db	3103	GAGGTCGTGTTCAATTTATTCACATCCACCGGACACAGATCCAGTCCAGGGCTGAGGCTC	316

QY 905 GCCAACATTTCTATGGTCCGCTTCTTATATGATCTATGATCATATATGCTCTGCGCTTC 964
|||||
Db 1210 GCCAACATTTCCATGGGTGCGACCTTTCTGCTATCTATGATCATATATGCTCTGCGCATTC 1269
QY 965 TGGTATGGACCTCCTTGGTCTCTCCAGTGAATATTCATATGGACAAGTACTGCTCAGCTGC 1024
|||||
Db 1270 TGGTATGGACCTCCTTGGTCTCTCAAAAGAAATACACTATTTGGACAAGTCTCAGCTGC 1329
QY 1025 TTCTTTCTGATTAATATGGGGCTTTTATGATTTGACAGGCGATCCCAAGCATTTGAAGCA 1084
|||||
Db 1330 TTTTCTTTCTGATTAATATGGAGCATTTCACTGTTGGCGGATCTCCAAATATTTGAAGCC 1389
QY 1085 TTTGCAACGCAAGAGGAGCAGCTTATGAATCTTCAAGATTAATGCAATTAACCAAGC 1144
|||||
Db 1390 TTGCGCAATGTAGAGAGCAGCTTATGAGCTTCAAGTATTAATGATTAATTAAGCCAGT 1449
QY 1145 ATTGACAGCTATTCGAAGAGTGGACATTAACCAATTAATTAATTAAGGAAATTTGGAATTC 1204
|||||
Db 1450 ATAGACAGCTTCTCAAAAGAGTGGGCACAAACCGACAAACATACAAAGGAAATTTGGAATTC 1509
QY 1205 AAAAATGTTCACTTCACTTACCTTCTGCAAAAGAAAGTAAAGCTTAAGAGGCTCAAC 1264
|||||
Db 1510 AAAAATATTTCACTTCACTTACCTTCTGCAAAAGAAAGCTTCAAGATCTTGAAGGCTCAAC 1569
QY 1265 CTGAAGGTTTCAGAGTGGGCGACAGTGGCGCTGGTTGGCAACAGTGGCTCGGGAGAGAG 1324
|||||
Db 1570 CTGAAGGTTGAAGAGGCGGACAGCGTAGCCCTGGTTGGCAACAGTGGCTCGGGAGAGAG 1629
QY 1325 AGCAGCGTGCAGCTGATTCAGAGCTCTATNAGCCCAACATGCGCATGCTCTATTTGAT 1384
|||||
Db 1630 ACACTGTCCAGCTGTCTGACAGAGGCTCTACAGCCCATTAAGGCGCAGTACATATCGAC 1689
QY 1385 GGACGAGCATTTAGACCATTAATGTAAGCATCTTGGGAAATTTAGTGGTGGTGAAT 1444
|||||
Db 1690 GGACGAGCATTCAGACCATCAATGTAAGCATCTTGGGAAATTTAGTGGTGGTGAAT 1749
QY 1445 CAGAGAGCTGTGTGTTGGCCACACAGATAGCTGAAAAAATTCGCTATGCGCGCGAAAT 1504
|||||
Db 1750 CAGAGAGCCGTGTGTGTTGGCCACACCAATTCGCGAAATTCGCTATGCGCGAGAAAC 1809
QY 1505 GTCAACATGATGATGATGGAAGGCTTAAAGAAAGCAATGCCATATATTTTATCATNG 1564
|||||
Db 1810 GTCAACATGATGATGATGGAAGGCTTAAAGAAAGCAATGCCATATATTTTATCATNG 1869
QY 1565 AAATCTACCTTAATTAATTTGACACTGTGTGAGAGAGAGAGGGGCCAGTGAATGATGA 1624
|||||
Db 1870 AAATCTGCCCAAAATTTGACACCTGTGTGAGAGAGAGGGGCCAGTGAATGATGA 1929
QY 1625 CAGAAACAGAGAAATCGGCATTTGCTGGGCGCTGTGCGCAACCCCAAGATTTCTGCTG 1684
|||||
Db 1930 CAGAAACAGAGAAATCGGCATTTGCTGGGCGCTGTGCGCAACCCCAAGATTTCTGCTG 1989
QY 1685 GATGAGCAACGTACAGCTGTGACACTGGAAGTGAAGAGAGAGGCGCCAGTGAATGATGA 1744
|||||
Db 1990 GATGAGGCGCACGTACAGCTGTGACACAGAAAGCGGCTGTGCTCAGGCGCGCTGTGAT 2049
QY 1745 AAGGCCAGAAAAGCGCGGACGTACCATTTGTGATGATGCTCATGTTGTCTCAAGTTCGTAAT 1804
|||||
Db 2050 AAGGCTGAGAAAGCGCGGACCATTTGTGATGATGCTCATGCGCTGTCTCAAGTTCGTAAT 2109
QY 1805 GCCGATGCTATTTGCTGTTTGTGATGAGATCTTGGGAAAGAAAGAAATCATGATGA 1864
|||||
Db 2110 GCGATGCTATTTGCTGTTTGTGATGAGATCTTGGGAAAGAAAGAAATCATGATGA 2169
QY 1865 CTGATGAAGAGAGGAGCTTACTTCAAACTGTACATGACAGACGACGAGAAATGA 1924
|||||
Db 2170 CTGATGAGAGAGAGGAGCTTACTTCAAACTGTACATGACAGACGACGAGAAATGA 2229
QY 1925 ATTGATTTGAAAATGCCATGCTGTAATCCAAATGAAAATGAAATGATGCTTTGGAATGCT 1984
|||||
Db 2230 ATTGATTTGAAAATGAACTGTTGTAATCTAAAGAYGGAATTTGTAATGTCGATGCTCT 2289
QY 1985 CCAAAAGATTCAGGCTCAGTTTAATAAAAGAAAGATCACTCGCAGAGAGTATCATGCA 2044
|||||
Db 2290 TCAAAAGATTCAGGATCCAGTCTAATAGAAAGATCACTCCGAAAGATCCGAGG 2349
QY 2045 CCACAGGCGCAAGACAGAAAGCTTGTACAAAAGAGGACTGAATGAGATGATCTCCA 2104
|||||
Db 2350 CCACATGATCAAAAGCGGGAAGCTTAGACACAAAGAGGCTCGATGAGAGAGCTCTCCA 2409
QY 2105 GTTTCCTCTGAGAGATTCCTGAGCTGAAGCTGAAGCTCAAGTGAATGCTTTTGGTGGT 2164
|||||
Db 2410 GCTTCCTTTGGCGATCCTGAAGTTGAATCACTGAATGCTTTATTTGGTGGTGGT 2469
QY 2165 AATATTTGGCTATTAATTAACGAGGCTGCAACGACGATTTTCAATTAATTTCAAG 2224
|||||
Db 2470 GATTTTGGCTAATTAATTAAGAGGCTTGCACAGATTTCTCCATTAATTTTCAAG 2529
QY 2225 ATTATGAGATCTTACCCGAGATGAGGATCTGAAACAAAGACAGAAATTAATGATG 2284
|||||
Db 2530 GTTGTAGGAGGTTTTCACAAAATATGACACCCTTAATTCAGCGGCGAAGACGCACTTG 2589
QY 2285 TTTTCTGATTTGTTCTAGTCTTGAATTAATTTCTTTATTAATTTTCTCCAGGCG 2344
|||||
Db 2590 TTTTCTTTATTTGTTCTGATCTTGGATCATCTTTCAATTAAGTTTCTCCAGGCG 2649
QY 2345 TTCATATTTGGCAAAAGCTGGGAGATCTCTACATGAGGCGCTTGATACATGTTTTCA 2404
|||||
Db 2650 TTCATATTTGGCAAAAGCTGGAGAGATCTCTACAGAGCTCCGATACATGATGTTCAAA 2709
QY 2405 TCATGCTGAGACAGATGTCAGCTGTTGATGATGACCTTAATAACACATGAGGATG 2464
|||||
Db 2710 TCATGCTGAGACAGATGTCAGCTGTTGATGATGATGATGATGATGATGATGATGATG 2769
QY 2465 ACAACGAGCTTGGCAATGATGCGGCTCAAGTTAAAGGGCTATGATGTTCAAGCTTGT 2524
|||||
Db 2770 ACAACGAGGCTTGGCAATGATGCGGCTCAAGTTAAAGGGCTATGATGTTCAAGCTTGT 2829
QY 2525 GTCAATACCCGAATATAGAAATCTTGGGAGAGGATTAATTAATCTTAATCTATGCT 2584
|||||
Db 2830 GTTATTAACCAAGATATAGAAATCTTGGGAGAGGATTAATTAATCTTAATCTATGCT 2889
QY 2585 TGCAATTAACACTTTTACTCTTAGCAATTTGATACCATTAATGATGATGATGATGAT 2644
|||||
Db 2890 TGCAATTAACACTTTTACTCTTAGCAATTTGATACCATTAATGATGATGATGATGAT 2949
QY 2645 GAAATGAAGATGTTGCTGAGCAAGCAAGCAAGCAAGCAAGCAAGCAAGCAAGCAAGCA 2704
|||||
Db 2950 GAAATGAAGATGTTGCTGAGCAAGCAAGCAAGCAAGCAAGCAAGCAAGCAAGCAAGCA 3009
QY 2705 AAGATGCTACAGAAAGCAAGCAAGCAAGCAAGCAAGCAAGCAAGCAAGCAAGCAAGCA 2764
|||||
Db 3010 AAGATGCTACAGAAAGCAAGCAAGCAAGCAAGCAAGCAAGCAAGCAAGCAAGCAAGCA 3069
QY 2765 AAGTTGATTAATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 2824
|||||
Db 3070 AAGTTGATTAATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 3129
QY 2825 GCACATATCTTGGGGTCTCATTTCTATACCCAGGCAAGTATATTTTCTATGCT 2884
|||||
Db 3130 GCGCAGCTCTTGGGATCTCTTCTCTTCTCAGCCAGGCGCATGATTAATTTCTCTATGCT 3189
QY 2885 GCGTGTTCGCGTTGGTGGTCTTCTTGTGGCAATGATGATGATGATGATGATGATG 2944
|||||
Db 3190 GCGTGTTCGCGTTGGTGGTCTTCTTGTGGCAAGTATGATGATGATGATGATGATG 3249
QY 2945 CTTTGTTATTTCTGAGCTTCTTCTTGTGGTATGATGATGATGATGATGATGATG 3004
|||||
Db 3250 CTTTGTTATTTCTGAGCTTCTTCTTGTGGTATGATGATGATGATGATGATGATG 3309
QY 3005 GCTCTGATATGCCAAAGCCAAATATACAGAGCCGACGATCATGATGATGATGATG 3064
|||||
Db 3310 GCTCTGATATGCCAAAGCCAAATATACAGAGCCGACGATCATGATGATGATGATG 3369
QY 3065 AGCCCTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 3124
|||||

Db 3370 ATCCCTGAGATTGACAGTACAGCAGGAGGCTTGAAGCCTAATATGTGGAAGAAAT 3429
Oy 3125 GTGACATTTAATGAGGTCTGTTCACATATCCCACTGGACACAGCATCCCGCTCCAG 3184
Db 3430 GTGAATTTAAATGAGATGTTCACATATCCCACTGGACACAGCATCCCGCTCCAG 3489
Oy 3185 GGGCTGAGGCTCGAGTGAAGAGGCGCAGACGCTGGCCCTCTAGGTAGCAGTGCCTG 3244
Db 3490 GGGCTGAGGCTCGAGTGAAGAGGCGCAGACGCTGGCCCTCTAGGTAGCAGTGCCTG 3549
Oy 3245 GGGAGAGCAGCATGTTTACGCTCTAGAGCGCTTCTATGACCCCTGGCTGGTTGAGTG 3304
Db 3550 GGGAGAGCAGCATGTTTACGCTCTAGAGCGCTTCTATGACCCCTGGCTGGTTGAGTG 3609
Oy 3305 CTAATGTAGTGAAGAGATTAAGACCTGATGTCAGTGGCTCCGAGCACACCTGGGC 3364
Db 3610 TTTCTAGATGGCAAAATAAAGCACTCAATGTCCAGTGGCTCCGAGCACACCTGGGC 3669
Oy 3365 ATCGTGTCTCAGGAGCCCATCTGTTTACGCTGACAGCATTCGCCGAGACATTCCTATGGA 3424
Db 3670 ATTGTGTCCAGGAGCCCATCTGTTTACGCTGACAGCATTCGCCGAGACATTCCTATGGA 3729
Oy 3425 GAGAAAGCCGGCTCTACATGAGATTAAGAGATTAAGAGGCGCAGAGGAGGAGGAGGAG 3484
Db 3730 GAGAAAGCCGGCTCTACATGAGATTAAGAGATTAAGAGGCGCAGAGGAGGAGGAGGAG 3789
Oy 3485 CACCACTTCATCGAGACCTCCCTGAGAAATCAACACACAGAGTGAAGAGAGAGAGAG 3544
Db 3790 CACCACTTCATCGAGACCTCCCTGAGAAATCAACACACAGAGTGAAGAGAGAGAGAG 3849
Oy 3545 CAGCTCTGTGTGGCCAGAAAGCGCATTCGTCATGAGTGGCGGCTGTTGTTGAGACAGCT 3604
Db 3850 CAGCTCTGTGTGGCCAGAAAGCGCATTCGTCATGAGTGGCGGCTGTTGTTGAGACAGCT 3909
Oy 3605 CATATTTTGTGTTGATGAGCTACATGAGCTGTGATGATGAGAGAGAGAGAGAGAGTGC 3664
Db 3910 CATATTTTGTGTTGATGAGCTACATGAGCTGTGATGATGAGAGAGAGAGAGAGAGTGC 3969
Oy 3665 CAAGAAAGCCCTGAGCAAGAGCGCAGAGAGCGCAGCTGATGTGATGCGCCACCGCTTG 3724
Db 3970 CAAGAAAGCCCTGAGCAAGAGCGCAGAGAGCGCAGCTGATGTGATGCGCCACCGCTTG 4029
Oy 3725 TCCACCATTCAGATGAGATTAATGATGTTTCAATGATGCAAGAGAGAGAT 3784
Db 4030 TCCACCATTCAGATGAGATTAATGATGTTTCAATGATGCAAGAGAGAGAT 4089
Oy 3785 GGCACATCAAGCAGCTGTGGCCAGAAAGCAGCATTTTTCATGCTGCTGCGAG 3844
Db 4090 GGCACATCAAGCAGCTGTGGCCAGAAAGCAGCATTTTTCATGCTGCTGCGAG 4149
Oy 3845 GCTGAGCAAGAGCGCTAGTACGTGGCCATGAGCTGTTAATTTTAAATTTAATTT 3904
Db 4150 GCTGAGCAAGAGCGCTAGTACGTGGCCATGAGCTGTTAATTTTAAATTTAATTT 4208
Oy 3905 GTGTTAAATGAGCATTATCAAGATTAAGAGTGAAGCATTACTGGAAG 3957
Db 4209 GTATTAATTAATGACACTATCAAGTCAAAAGAGAAACATTACTAGATA 4261

RESULT 9
US-09-769-097-1
; Sequence 1, Application US/09769097
; Patent No. US20020055128A1
; GENERAL INFORMATION:
; APPLICANT: Kimberly Anne Brun
; APPLICANT: Richard James Chenery
; APPLICANT: Harma Eilens
; APPLICANT: John Anthony Feld
; APPLICANT: Lin Yue
; TITLE OF INVENTION: POLYNUCLEOTIDE AND POLYPEPTIDE SEQUENCES
; TITLE OF INVENTION: ENCODING RAT MDRIA AND SCREENING METHODS THEREOF
; FILE REFERENCE: GP-5009-C2
; CURRENT APPLICATION NUMBER: US/09/769,097

; CURRENT FILING DATE: 2001-01-25
; PRIOR APPLICATION NUMBER: 09/208,809
; PRIOR FILING DATE: 1998-12-09
; PRIOR APPLICATION NUMBER: 09/156,800
; PRIOR FILING DATE: 1999-09-17
; PRIOR APPLICATION NUMBER: US99/20770
; PRIOR FILING DATE: 1999-09-10
; NUMBER OF SEQ ID NOS: 4
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 1
; LENGTH: 4369
; TYPE: DNA
; ORGANISM: RATTUS RATTUS
US-09-769-097-1

Query Match 68.2%; Score 2917; DB 10; Length 4369;
Best Local Similarity 84.3%; Pred. No. 0;
Matches 3332; Conservative 0; Mismatches 590; Indels 31; Gaps 3;

Oy 5 CCGGAGGTGGGATGATCTGTAAGAGGCGCTAAGGGAGTGCAGAGAACTTCTGG 64
Db 340 CGTAGGTGCTGATGAGCTCGAAGAAACCTTAACGGAAGACAGACAAAGAACTTCTCA 399
Oy 65 AAAATGGCAAAAAAGTAAATAATGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 124
Db 400 AAGATGGCAAAAAAGTAAATAATGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 456
Oy 125 TTTGCAATTTTGGCTATTCATTAATGAGTGTGATGATGATGATGATGATGATGATG 184
Db 457 CTCACATGTTTGGCTATTCATTAATGAGTGTGATGATGATGATGATGATGATGATG 516
Oy 185 GCTGCCATCATCATGAGTGCCTCCCTCATGATGATGATGATGATGATGATGATGATG 244
Db 517 GCGGCATTCATCATGAGTGCCTCCCTCATGATGATGATGATGATGATGATGATGATG 576
Oy 245 GATAGCTTTGCAATTCAGAGATTTCAAGAAACAACTTTCCAGTATTAATGA 304
Db 577 GATAGCTTTGCAATTCAGAGATTTCAAGAAACAACTTTCCAGTATTAATGA 617
Oy 305 AGTATTAGCAACATTCACACATTCATCAACATTCATCAACATTCATCAACATTCATCA 364
Db 618 -----CTACAACTCTACACATTCATCAACATTCATCAACATTCATCAACATTCATCA 669
Oy 365 GCGTATTATTACGTGGATGCGTGGCTGGCTGGCTGGCTGGCTGGCTGGCTGGCTGGCT 424
Db 670 GCGTATTATTACGTGGATGCGTGGCTGGCTGGCTGGCTGGCTGGCTGGCTGGCTGGCT 729
Oy 425 TTTGTTGCTGCTGACAGAGAGAGATGATCAAAATTAAGAAACAAATTTTTCATGCT 484
Db 730 CTTTGGTCTGCTGACAGAGAGAGATGATCAAAATTAAGAAACAAATTTTTCATGCT 789
Oy 485 ATCATGAGAGAGATGCGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGG 544
Db 790 ATCATGAGAGAGATGCGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGG 849
Oy 545 CTCACAGAGAGTGTCCAAATCAATGAAGAGATGGGAGACAAATTTGGAATGTTCTTT 604
Db 850 CTCACAGAGAGTGTCCAAATCAATGAAGAGATGGGAGACAAATTTGGAATGTTCTTT 909
Oy 605 CACTCAATAGCAATTTTACCGGTTTATAGTGGGTTTACAGCTGTTGAGAGAG 664
Db 910 CAGCAATAGCAATTTTACCGGTTTATAGTGGGTTTACAGCTGTTGAGAGAG 969
Oy 665 ACCCTTGATTTGGCAGTACGCGGTCTGGAGCTTTCAGCCGCACTGCGGCAAG 724
Db 970 ACTCTTGATTTGGCAGTACGCGGTCTGGAGCTTTCAGCCGCACTGCGGCAAG 1029
Oy 725 ATACATCTTATTTAGTGAATTAAGAGCTTGGCTATGCAAAAGCTGAGCAGTACT 784
Db 1030 ATATGCTTATTTAGTGAATTAAGAGCTTGGCTATGCAAAAGCTGAGCAGTACT 1089
Oy 785 GAGCAAGTCTTAGCAGCAATCAGACTGATGATGCTTGGAGAGAGAGAGAGAGACTT 844


```

FILE REFERENCE: 1340-1-02C1C2
CURRENT APPLICATION NUMBER: US/09/866, 866A
CURRENT FILING DATE: 2001-08-30
PRIORITY APPLICATION NUMBER: 09/584, 586
PRIORITY FILING DATE: 2000-05-31
PRIORITY APPLICATION NUMBER: PCT/US99/11825
PRIORITY FILING DATE: 1999-05-27
PRIORITY APPLICATION NUMBER: 60/086, 988
PRIORITY FILING DATE: 1998-05-28
NUMBER OF SEQ ID NOS: 27
SOFTWARE: Patentin version 3.0
SEQ ID NO 5
LENGTH: 4189
TYPE: DNA
ORGANISM: Mus musculus
US-09-866-866A-5

Query Match      63.5%;      Score 2716.6;      DB 10;      Length 4189;
Best Local Similarity  81.7%;      Pred. No. 0;
Matches 3215;      Conservative  0;      Mismatches  694;      Indels  25;      Gaps  6

```

[illegible]

QY	794	TTTGACAGCATCAGACATGTGATTGGCTTTGGAGGAGCAAAAGAAAGAACTTGGAAAGTAC	855
Db	769	TTAGGAGCCATCAGAACTGTGATTGGCTTTGGAGGAGCAACACAGAGAACTTGAAGGATAC	828
QY	854	AAACAAAATTTGAAGAAAGCTTAAGGAAATTGGATTAAGAAAGCATACCGCCACATTT	913
Db	829	AAATAAAAATTTGAAGAAAGCTAAAAATGTTGGCATTAAGAAAGCATATCACAGCCAGATTT	888
QY	914	TCATTTGTGCGCGCTTCTCTATTGATATATGATCATATATGCTCTGGCTTTTGGTATGGG	973
Db	889	TCGATAGAGCATTCGCTACCTGTTGGTCTATATGATCATATATGACATCTTCGTATATGGG	948
QY	974	ACCTCCCTGGTCCCTCTCAGATGAAATTTCTATTGGACAAGTACTACATGCTCTTCTTTCTC	103
Db	949	ACATCTCTGGTCCCTCTCAATTAATTTCTATTGGAGAAAGTGTCTATGCTTCTTCTCTCT	100
QY	1034	GTATTTAATTTGGGGCTTTTAGTATTATTTGACAGGCAATCCCAACGATTTGAACATTTGCAAC	109
Db	1009	ATTTTGTGTGGGACTTTTATATATATGACACTTGGCCCAACATATAGAAAGCTTTGCAAC	106
QY	1094	GGAAGAGAGAGAGCTTTTGAATTTCTCAAGATATTTGCAATTAACCAACGATTTGACAG	115
Db	1069	GACAGAGGGGAGGCTTTGAAATCTTCAAGATATTTGATTAACAGCCAAACATTTGACAC	112
QY	1154	TATTCGAAGAGTGACATTAACACAGATTAATTTAAGGAAATTTGGAAATTCAAAAATGTT	121
Db	1129	TTCTCAACAAAGGGCTTACAAACACACAGTATATATGGAAACTTAGAGTTTAAAAATGTT	118
QY	1214	CACCTTCAGTTACCTCTCTGAAAAGAAAGTATTAAGTATCTTTAAAGGCTCTCAACCTGAAGTT	127
Db	1189	CACCTTCAACTACCCATTCAGAGAAAGCCAACTTCAGATCTTGAAGGGCCCTCAATCTGAAGGTG	124
QY	1274	CAGAGTGGGACAGACATGTGGCTGGTGTGGAAAGTGGCTGGGGAAGGACACACACCTG	133
Db	1249	AAGAGCGGACAGACGGGTGGCTTGGTGGCAACAGTGGCTGTGGAAAGGACCAACTCTC	130
QY	1334	CAGCTGATGAGAGGCTCTATATGACCCACAGATGGCAATGGCTGTATATATGAGACAGAGC	139
Db	1309	CAGCTATATGAGAGGCTCTACGACCCCTGGAGGGGCGTGGTCAGTATCGAGGACAAACAC	136
QY	1394	ATTAGGACCATAAATATTAAGCATCTTCTGGGAAATTAACGTGGTGGTGAAGTACGAGACT	145
Db	1369	ATCAGAACCATATATGTAGGTATCTAGGGAATCATATGGTGGTGAAGTACAGAACT	142
QY	1454	GTGTGTTTGGCACACGATAGCTGTGAAGAACTTGGTATGGCCCGGAAATATGTACCATG	151
Db	1429	GTGTGTTTGGCACACGATAGCTGTGAAGAACTTGGTATGGCCCGGAAATATGTACCATG	148
QY	1514	GATGAGATTGAAAGAGCTGTTAAGGAAAGCCAAATGCTATGATTTATATGAATCTACT	157
Db	1489	GATGAGATTGAAAGAGCTGTTCAGAGAGAGCCAAATGCTATGATTTATATGAATCTCTCC	154
QY	1574	AATTAATTTGACACTGTGGTTTGAAGAGAGAGGGGCGCAGTGAAGTGGTGGAGCAAGAAACG	163
Db	1549	CACCAATTTGACACCTGTGGTTGGTGAAGAGAGGGGCGCAGTGAAGTGGGGAACAAACAG	160
QY	1634	AGAAATCGCATTTGTGGGGCCCGGTGTCCCAACCCCAAGATTTCTTGTGTGATGAGGCA	169
Db	1609	AGAAATCGCATTTGTGGGGCCCGGTGTCCCAACCCCAAGATCTTTTGTGTGAGCAGAGCC	166
QY	1694	ACGTAGCTGTGGACACTGAAAGTGAAGAGTGAAGTTCAGTGGGCGCTGGATTAAGGGCACA	173
Db	1669	ACCTCAGCCCTGGATACAGAAAGTGAAGTGTGGTGGCAGCGCCGACCTGGATTAAGGCTTAA	172
QY	1754	AAAGGCGGAGCTACCATTTGTATGATGCTCATGCTTTGTCTACAGTTGCTAATGCCATGTC	181
Db	1729	GAAAGCGGAGCAACATTTGTATGATGCTCATGCTTTGTCTACAGTTGCTAATGCCATGTC	178
QY	1814	ATTCTGTGTTTATGATGAGATCATTTGTGAGAAAGAAATCATGATGACATCATGAGAA	187
Db	1789	ATTCTGTGTTTATGATGATGATGATCATTTGTGAGAAAGAAATCATGATGACCTATGAGA	184
QY	1874	GAGAAAGGCACTTACTTCAAACTTCTCAATATGACAGCAAGAGAAATGAATTAAGTTA	193

D	1849	GAAGAAGCATTTTCTTCAAACTTGTATGACACAGACTAGAGAAATGAAATGGAACCA	1901
Q	1934	GAAAATGCCACTGGTGAATCCAAAAGTGAAGATGAGCTTGGAAATGTCCTCAAAAAGT	199
D	1909	GGAATATATCTTATGATGCCAGAGTGAACACTGATGCTTCTGAACTGACTTCAGAAAGA	196
Q	1994	TCAGGTCGCAAGTTTAATAAAAAAGATCAACTCGCAGAGATATACATGACCACAGAAGC	205
D	1969	TCCAAATACCTTTAAT--AAGGAATCAATTTACAGAAAGTGCTCCAGAAAGCAAGC	202
Q	2054	CAAGACGAAGAAGCTTGATACAAAAGAGGACTGGAATGAAATGATGACTCAAGTTTCCTC	211
D	2026	CAAGAGAGAAAGACTATGATGTAAGAAAGGCGTGTGATGAAGAATGCTCTGTGGTTTCCTT	208
Q	2114	TGAGAGATTCGAGCTGAACTCAACTGAATGAGCCCTATTTTGTGTGTATATTTTGT	217
D	2086	TGGCGAGATCTTAATATCTTAATCTAAGTGAATGCCCTATTTACTGTGTGGCTACTTTCG	214
Q	2174	GCTATTTAAAGGAGGAGCGCTGCAACAGACTTTTCAATATATTTTCAAGATATAGG	223
D	2146	GCTGTTATTAATGGGTGCATACAAACAGTGTTCGCAATATGATTTTCAAGATATGATGG	220
Q	2234	ATCTTACCCGAGATGAGATCTTGAACAAACAGCAGATAGTAACTGTTTTCGTA	229
D	2206	GTTTTTTCAAGAGATGATGACCATGAACCTAAACAGAGATTTGATTTTTCCTCG	226
Q	2294	TGTGTTAGTCTTGGAAATATTTCTTTATTTATCTTTTCCAGAGGCTTCAATTT	235
D	2266	TTCCTTGTGTTATGAGGCTGATTTCTTTGTTCATATTTCTTTCAGGGCTTCCACTTT	232
Q	2354	GGCAAAGCTGGGAGATCCTCACTAAGCGGCTTGATATACATGTTTCAGATCCATCTG	241
D	2326	GGCAAAGCGGAGAGATCTCCACCAAGGAGTCCGATATACATGTTTCAATTCATGCTG	238
Q	2414	AGACAGATGTCAGTGGTTGATGACCCCTAAAAACACCACTGGAGCATTTGACAACAG	247
D	2386	AGACAGATATATACCTGTTGATGACCATTAAGAACAGCACTGGCTCACAGCACCAAG	244
Q	2414	CTTGCAATGATGCGGCTCAAGTTAAAGGGGCTATAGTTTCCAGGCGTGTGTCATTACC	253
D	2446	CTCCCAAGTATGCTCTAGTGTTAAAGGGCGCATGGGCGCCAGGCTGTGATGTTACC	250
Q	2534	CAGATATGAGAAATCTTGGGACAGGCAATTAATATCTTAATCTATGTTGGCAATTA	259
D	2506	CAGATATGAGAAACCTCGGGAAGAGATCATCTCTCTTATGTCATGCTGGACGCTG	256
Q	2594	ACACTTTTACTCTTAGCAATTTAGCCATATTTGCAATTTGCAATGAGAGTGTGAATGAA	263
D	2566	ACACTTACTTGTAGTAATTAATACCGCTCATTTGATTTGGCGGAATTAATTAATGAAG	262
Q	2654	ATGTTGCTGACACAGCACTGAAGAATTAAGAAAGAGCTAGAAGGAGCTGGGAAGTTGCT	271
D	2626	CTGTTGTCTGGCCCAACCTTGAAGGACAGAAACAGCTTGAAATCTCTGGAAAGTTGCT	268
Q	2714	ACAGAAGCATGCAAAACTTCGAACTGTGTTTCTTTGACTCGGAGCAGAAAGTTTGA	277
D	2686	ACAGAAGCAATTAAGAACTTCGCACTATTTGCTCTTTGACTCGGAGCAGAAAGTTTGA	274
Q	2774	TACATATGACAGAGTTTGCAGTACCATACAGAAACTCTTTGAGGAAGACACATC	283
D	2746	ACCATGTATGCCCAAGCTTGAAGTACCATTAAGAAATGCGATGAAGAAAGCACAGTG	280
Q	2834	TTTCGGGCTTCAATTTTCAATCCAGCAATGATGATATTTTCTTATGCTGAGCTGTTTC	289
D	2806	TTTGGGATACCTTCTCTTCAACCCAGGCATGATGATTTTCTTATGCTGTTTC	286
Q	2894	CGGTTTGTGCTCTTGGTGGCAAAATAGTTCATGAACTTTCAGATGCTCTTTTGGTA	293
D	2866	CGGTTGCTGCTACTTGTGTGGCAACAACATCATGACTTTTGAAGAAATGATATGGTGGTA	292
Q	2954	TTTCAGTATTTGCTTGTGTGCTGACATGGCAGTGGGCAAGGTCAATTCATTTGCTCGTAC	301

```

Db 2926 TTTTCTGCTGTTGTCTTTGGTSCATGAGCAGCTGGGAATTAAGTTATTTGCTCTGAC 2985
OY 3014 TATGCCAAAGCCAAAGTATACAGACCCGACGTCATGATCATTTGAAAAAGCCCTCTG 3073
Db 2986 TATGCCAAAGCCAAAGTATACAGATCTATATCATCAGATCATTTGAAAAAGCCCTGAG 3045
OY 3074 ATTGACAGCTACAGCCCTCAGGCTTCAAGCCAAATACGTTGGAAGAAATGTGACATTT 3133
Db 3046 ATTGACAGCTACAGCAGCAGAGGGCTTGAAGCTTACTCTGTAGGAAGAAATGTAAATTT 3105
OY 3134 AATGAGCTCTGTTCAACTATCCACTGACAGACATCCCGCTCTCCAGGGCTGAGC 3193
Db 3106 AATGAGCTCTGTTCAACTATCCACTGACAGACATCCCGCTCTCCAGGGCTGAGC 3165
OY 3194 CTGAGGTGAAGAAGGCGCAGACGCTGGCCCTGTAGTAGACAGTGCTGTGGAAGAGC 3253
Db 3166 CTGAGGTGAAGAAGGCGCAGACGCTGGCCCTGTGGAAGAGTGCTGTGGAAGAGC 3225
OY 3254 ACACTTGTCACTCTTAGAGCGCTTCTATGACCCCTTGGCTGCTCACTGCTAAATGAT 3313
Db 3226 ACACTTGTCACTCTTAGAGCGCTTCTATGACCCCTTGGCTGCTCACTGCTTCTAGAT 3285
OY 3314 GGCAGAGATTAAGACACTGAAATGTCAGTGGCTCCGAGCACACTGGGCATCTGTCT 3373
Db 3286 GGCAGAGATTAAGACACTGAAATGTCAGTGGCTCCGAGCACACTGGGCATCTGTCT 3345
OY 3374 CAGAGCCCATCTCTGTTGACTGACATGTCGAGAGAACTGCTATGAGACACACAGC 3433
Db 3346 CAGAGCCCATCTCTGTTGACTGACATGTCGAGAGAACTGCTATGAGAGCAACAG 3405
OY 3434 CGGGTGTATCATAGAGATTAATGAGAGCAGCAGCCAGAGGCGCAATACACCATTC 3493
Db 3406 CGGGTGTATCATAGAGATTAATGAGAGCAGCAGCCAGAGGCGCAATACACCATTC 3465
OY 3494 ATCGAGACACTCTCTGAGAAATACACACAGAGTAGAGAGCAAGAGGAGGCTCTCT 3553
Db 3466 ATCGAGACACTCTCTGAGAAATACACACAGAGTAGAGAGCAAGAGGAGGCTCTCT 3525
OY 3554 GGTGGCCAGAAACAGCAGATTCCTATGCTGCGCTCTTGTAGACAGCTCATATTTTG 3613
Db 3526 GGTGGCCAGAAACAGCAGATTCCTATGCTGCGCTCTTGTAGACAGCTCATATTTTA 3585
OY 3614 CTCTTGTAGTAAGTACATCACTCTGATACAGAAAGTGAAGGTTGTCCAAAGAGCC 3673
Db 3586 CTCTGAGAGAGCAACATCACTCTGATACAGAAAGTGAAGGTTGTCCAAAGAGCG 3645
OY 3674 CTGAGAGAGAGCAAGAGGCGCAGCTGATGATGCGCCACCGCTTTCACATC 3733
Db 3646 CTGAGAGAGAGCAAGAGGCGCAGCTGATGATGCGCCACCGCTTTCACATC 3705
OY 3734 CAGAAATGCAATTAATAGTGTGTTTCAGAAATGCAAGAGTGAAGAGCATGACACAT 3793
Db 3706 CAGAAATGCAATTAATAGTGTGTTTCAGAAATGCAAGAGTGAAGAGCATGACACAC 3765
OY 3794 CAACAGCTGTGGCCAGAAAGGCACTTATTTTCCATGTCAGTGTCCAGGCTGAGCA 3853
Db 3766 CAGAGCTGTGGCCAGAAAGGCACTTATTTTCCATGTCAGTGTCCAGGCTGAGCA 3819
OY 3854 AAGGCTATGCACTGTGGCCATATGAGCTGTAAATATTTTAAATATTTGTTTAA 3913
Db 3820 AAGGCTATGCACTGTGGCCATATGAGCTGTAAATATTTTAAATATTTGTTTAA 3878
OY 3914 CATGCACTTAAATCAAGTTAAAGTGAAGGACT 3947
Db 3879 CATGCACTTAAATCAAGTTAAAGTGAAGGACT 3912

```

```

RESULT 11
US-09-917-800A-1424
; Sequence 1424, Application US/09917800A
; Patent No. US20020119462A1
; GENERAL INFORMATION:
; APPLICANT: Mendrick, Donna

```

```

; APPLICANT: Porter, Mark
; APPLICANT: Johnson, Kory
; APPLICANT: Castle, Arthur
; APPLICANT: Elashoff, Michael
; APPLICANT: Gene Logic, Inc.
; TITLE OF INVENTION: Molecular Toxicology Modeling
; FILE REFERENCE: 44921-5038-US
; CURRENT APPLICATION NUMBER: US/09/917,800A
; PRIOR FILING DATE: 2001-07-31
; PRIOR APPLICATION NUMBER: US 60/222,040
; PRIOR FILING DATE: 2000-07-31
; PRIOR APPLICATION NUMBER: US 60/222,880
; PRIOR FILING DATE: 2000-11-02
; PRIOR APPLICATION NUMBER: US 60/290,029
; PRIOR FILING DATE: 2001-05-11
; PRIOR APPLICATION NUMBER: US 60/290,645
; PRIOR FILING DATE: 2001-05-15
; PRIOR APPLICATION NUMBER: US 60/292,336
; PRIOR FILING DATE: 2001-05-22
; PRIOR APPLICATION NUMBER: US 60/295,798
; PRIOR FILING DATE: 2001-06-06
; PRIOR APPLICATION NUMBER: US 60/297,457
; PRIOR FILING DATE: 2001-06-13
; PRIOR APPLICATION NUMBER: US 60/298,884
; PRIOR FILING DATE: 2001-06-19
; PRIOR APPLICATION NUMBER: US 60/303,459
; PRIOR FILING DATE: 2001-07-09
; NUMBER OF SEQ ID NOS: 1740
; SOFTWARE: Patent Ver. 2.1
; SEQ ID NO 1424
; LENGTH: 4254
; TYPE: DNA
; ORGANISM: Rattus norvegicus
; FEATURE:
; OTHER INFORMATION: Genbank Accession No. US20020119462A1 M81855
; US-09-917-800A-1424

Query Match 62.6%; Score 2677.4; DB 10; Length 4254;
Best Local Similarity 81.4%; Pred. No. 0;
Matches 3212; Conservative 0; Mismatches 701; Indels 34; Gaps 8;

OY 5 GCGGAGTGGGATGATCTGTAAGAGGCGGTAAAGGAGTGCAGAGAAATCTCTG 64
Db 91 CGTGAGTCTGATGAGATTTGAAGAGGCGCTTAACGAGAGCAAGAACTCTCA 150
OY 65 AAATGGCAAAAAGTAAATGAAGAAAGAAAGAAAGAAAGAAAGAAAGTGCAGCAG 124
Db 151 AAGATGGCAAAAAGTAAATGAAGAAAGAAAGAAAGAAAGAAAGTGCAGCAGT 204
OY 125 TTGCAATGTTTCGATCAATTAATGATGATGATGATGATGATGATGATGATGATG 184
Db 205 TTGCGATGTTTCGATCAATTAATGATGATGATGATGATGATGATGATGATGATG 264
OY 185 GCTGCATATCCATGAGCTGACCTCTCTCATGATGCTGCTTGTGGAAGATGACA 244
Db 265 GCTGCATATCCATGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 324
OY 245 GATAGCTTTCAGAAATGAGAAATTTCAAGAAAGAAAGTTCAGTTATTAATGA 304
Db 325 GATAGTTT-----TACCCAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 378
OY 305 AGTATTAAGAAATACACATCAATTTTCATCAGC-----CATCGAGAGAGAGAGAGAG 361
Db 379 AGTAAATTAAGAAATACACATCAATTTTCATCAGC-----CATCGAGAGAGAGAGAG 438
OY 362 TATGCTATTAATTAAGAGTGGATGCTGCTGAGGCTGCTGCTGCTGCTTACATCA 421
Db 439 TATGCTATTAATTAAGAGGAGGATGCTGCTGAGGCTGCTGCTGCTTACATCAAG 498
OY 422 TCATTTGCTGCTGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 481
Db 499 TCATTTGCTGCTGAGAGGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 558

```

QY 482 GCTATCATCGACAGAGATTGGCTGTTTGAACGATGACGTTGGGAGCTTAACACC 541
DB 559 GCCATCATGAATCAGAGATATAGCGTGGTTTACGCGATGACGCTGGGGAGCTCAACACC 618
QY 542 CGGCTCACAGACGATGTCCTCAAAATCAATAGAGAAATGGCGCAAAATTTGAAATGTC 601
DB 619 GCTCTCACAGATGACGCTCCAAAATTAATAGCAAGATTTGGTGAACAACCTGGAAATGTTT 678
QY 602 TTTCACTCAATAGCAAACTTTTTCACCGGTTTATAGTGGGTTTACACGTGGTTGGAAG 661
DB 679 TTTGAGTCATAGACACTTTTTCACCGGTTTATTAATAGCAATTTAAAGTGGTTGGAAG 738
QY 662 CTAAACCTTGATTTGGCCATCAGCCCTGTTCTTGAGACTTACACCGCATCTGAGCA 721
DB 739 CTAAACCTTGATTTGGCCATCAGCCCTGTTCTTGAGACTTACACCGCATCTGAGCA 798
QY 722 AAGTACTATCTTACTTACTGATTAAGAACTCTTGCCCTATGCAAAAGCTGGAGCAGTA 781
DB 799 AAGTACTATCTTACTTACTGATTAAGAACTCTTGCCCTATGCAAAAGCTGGAGCAGTA 858
QY 782 GCTAAGAACTCTTATGACCAATCAGAACTGATTTGCTTTGGAGGACAAAAGAAAGA 841
DB 859 GCGCAAAAGCTTATGACCAATCAGAACTGATTTGCTTTGGAGGACAAAAGAAAGA 918
QY 842 CTGGAAGGATCAACAAAAATTTAGAAAGCTAAAGAAATTTGGATTAAGAAAGCTATTC 901
DB 919 CTGGAAGGATCAACAAAAATTTAGAAAGCTAAAGAAATTTGGATTAAGAAAGCTATTC 978
QY 902 AGCGCCAACTTCTATTTGCTCCGCTTCTATTTGATCTATGCAATGCTATGCTCTGCT 961
DB 979 AGCGCCAACTTCTATTTGCTCCGCTTCTATTTGATCTATGCAATGCTATGCTCTGCT 1038
QY 962 TTCTGATGAGCACTCTCTGCTCTCTCAGTGAATTTCTATTTGCAAGTACTACT 1021
DB 1039 TTCTGATGAGCACTCTCTGCTCTCTCAGTGAATTTCTATTTGCAAGTACTACT 1098
QY 1022 GTCTCTTTTCTGATTAATTTAGGCTTTTATTTGATGACAGGATCCCAAGCATTTGAA 1081
DB 1099 GTCTCTTTTCTGATTAATTTAGGCTTTTATTTGATGACAGGATCCCAAGCATTTGAA 1158
QY 1082 GCATTTGCAAGGCAAGAGAGAGAGAGCTTAAGAAATTTCAAGATTAATGCAATTAACCA 1141
DB 1159 GCATTTGCAAGGCAAGAGAGAGAGAGCTTAAGAAATTTCAAGATTAATGCAATTAACCA 1218
QY 1142 AGCATTTGACAGCTATTTGCAAGAGTGAATTAACAGATTAATTAAGGAAATTTTGAA 1201
DB 1219 AGCATTTGACAGCTATTTGCAAGAGTGAATTAACAGATTAATTAAGGAAATTTTGAA 1278
QY 1202 TTCAAAATTTTCACTTCACTTCACTTCACTTCACTTCACTTCACTTCACTTCACTTCA 1261
DB 1279 TTCAAAATTTTCACTTCACTTCACTTCACTTCACTTCACTTCACTTCACTTCACTTCA 1338
QY 1262 AACCTGAAGTTTCACTTCACTTCACTTCACTTCACTTCACTTCACTTCACTTCACTTCA 1321
DB 1339 AACCTGAAGTTTCACTTCACTTCACTTCACTTCACTTCACTTCACTTCACTTCACTTCA 1398
QY 1322 AGCAGGACCTTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGA 1381
DB 1399 AGCAGGACCTTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGA 1458
QY 1382 GATGAGACGAGCATTTAGGACCATTAATTAAGGATCTTTGGGAAATTTAGTGGTGTG 1441
DB 1459 GATGAGACGAGCATTTAGGACCATTAATTAAGGATCTTTGGGAAATTTAGTGGTGTG 1518
QY 1442 AGTCAGAGAGCTGTGTTGTTGCCAACAGATAGCTGAAGAAACATTCGCTATGAGCGGAA 1501
DB 1519 AGTCAGAGAGAGCTGTGTTGTTGCCAACAGATAGCTGAAGAAACATTCGCTATGAGCGGAA 1578
QY 1502 AATGTACCATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 1561
DB 1579 AATGTACCATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 1638
QY 1562 ATGAAACTTCTTAATTAATTTGACACTCTGTTGAGAGAGAGGCGCCAGCTGACGTGT 1621

DB 1639 ATGAAACTTCTTAATTAATTTGACACTCTGTTGAGAGAGAGGCGCCAGCTGACGTGT 1698
QY 1622 GGACAGAAACAGAGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 1681
DB 1699 GGACAGAAACAGAGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 1758
QY 1682 CTGATGATGAGCAAGCTCACTTGAAGCTGAGACATGAAAGTGAACAGTGGTGAAGTGGCC 1741
DB 1759 TTGATGATGAGCAAGCTCACTTGAAGCTGAGACATGAAAGTGAACAGTGGTGAAGTGGCC 1818
QY 1742 GATAAGGCAAGAAAGGCGGAGTCACTTGAAGCTGATGATGATGATGATGATGATGATGAT 1801
DB 1819 GATAAGGCAAGAAAGGCGGAGTCACTTGAAGCTGATGATGATGATGATGATGATGATGAT 1878
QY 1802 AATCCCATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 1861
DB 1879 AATCCCATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 1938
QY 1862 GAACTTGAAGAGAGAGGCACTTACTTCAAACTTGTCAATGCAATGCAAGAGAAAT 1921
DB 1939 GAACTTGAAGAGAGAGGCACTTACTTCAAACTTGTCAATGCAATGCAAGAGAAAT 1998
QY 1922 GAAATTTGATTTAGAAATGCACTGATGATGATGATGATGATGATGATGATGATGATG 1961
DB 1999 GAAATTTGATTTAGAAATGCACTGATGATGATGATGATGATGATGATGATGATGATG 2058
QY 1982 TTCCAAAGATTTCAAGGCTTCACTTAAATTAAGAGATCAACTGCGAGAGATATCAT 2041
DB 2059 TTCCAAAGATTTCAAGGCTTCACTTAAATTAAGAGATCAACTGCGAGAGATATCAT 2115
QY 2042 GCAACCAAGGCAAG 2101
DB 2116 GCAACCAAGGCAAG 2175
QY 2102 CCAAGTTCCCTTGAAGATTTCTGAAGCTGAATGATGATGATGATGATGATGATGATG 2161
DB 2176 CCAAGTTCCCTTGAAGATTTCTGAAGCTGAATGATGATGATGATGATGATGATGATG 2235
QY 2162 GGTATATTTTGTGATTAATTAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 2221
DB 2236 GGTATATTTTGTGATTAATTAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 2295
QY 2222 AGAATTAAGGATCTTATACCGAGATGAGAGATCTGAAACAAACAGAGATTAATAC 2281
DB 2296 AGAATTAAGGATCTTATACCGAGATGAGAGATCTGAAACAAACAGAGATTAATAC 2355
QY 2282 ATGTTTCTGATTAATTTCTAGTCTTGAATTAATTTCTTATTTTCACTTTTCTCCAG 2341
DB 2356 ATGTTTCTGATTAATTTCTAGTCTTGAATTAATTTCTTATTTTCACTTTTCTCCAG 2415
QY 2342 GCGTTCACTTTGGCAAGAGTGGGAGATCTCACTAAGCGGCTTGCATATGATGATGATG 2401
DB 2416 GCGTTCACTTTGGCAAGAGTGGGAGATCTCACTAAGCGGCTTGCATATGATGATGATG 2475
QY 2402 AGATTCATGCTGAGAGAGATGTCAGCTGTTGATGATGATGATGATGATGATGATGAT 2461
DB 2476 AGATTCATGCTGAGAGAGATGTCAGCTGTTGATGATGATGATGATGATGATGATGAT 2535
QY 2462 TTACAAACCAAGGCTTCCATGATGATGATGATGATGATGATGATGATGATGATGATG 2521
DB 2536 TTACAAACCAAGGCTTCCATGATGATGATGATGATGATGATGATGATGATGATGATG 2595
QY 2522 GCTGTCACTTACCAAGATTTATGATGATGATGATGATGATGATGATGATGATGATGAT 2575
DB 2596 GCTGTCACTTACCAAGATTTATGATGATGATGATGATGATGATGATGATGATGATGAT 2655
QY 2576 ATCTATGTTGGCAATTAACATTTTACTTATGATGATGATGATGATGATGATGATGATG 2635
DB 2656 ATCTATGTTGGCAATTTACTTATGATGATGATGATGATGATGATGATGATGATGATG 2715
QY 2636 GGAGTTGTTGAATGAATGATGATGATGATGATGATGATGATGATGATGATGATGATG 2695

Db 2716 GGAATTATTGAATGAACCTGTTGCTGTCAGCCTTGAAGACAAAGAAAGAGCTAGAG 2775
Qy 2696 GGAGCTGGAGAGATTGCTACAGAAAGCCATGAAAACCTTCCGAACTGTTGTTCTTACT 2755
Db 2776 AACTCTGGGAGATGCTACAGAAAGCAATGAAAACCTTCCGAACTGTTGTTCTTACT 2835
Qy 2756 CGGAGACAGATTGATACATGATATGACAGAGTTTGCAGATACATACAGAACTCT 2815
Db 2836 CGGAGACAGATTGATACATGATATGACAGAGTTTGCAGATACATACAGAACTCT 2895
Qy 2816 TTGAGAAAGACACATCTTCGGGCTCATTTTCTATCACCAGCAATGATGATTTT 2875
Db 2896 TTGAGAAAGACACATCTTCGGGCTCATTTTCTATCACCAGCAATGATGATTTT 2955
Qy 2876 TCCATGCTGCTGTTCCGGTTGGTGGCTACTTGGTGGCAATGATGATGATTT 2935
Db 2956 TCCATGCTGCTGTTCCGGTTGGTGGCTACTTGGTGGCAATGATGATGATTT 3015
Qy 2936 CAGAGATGCTTTTGGTATCTCAGCTATTTGCTTGGTGGCTAGTGAGGAGGCTC 2995
Db 3016 GAAATGTTATGTTGATTTTCTGCTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTT 3075
Qy 2996 AGTTCATTTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 3055
Db 3076 AGTTCATTTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 3135
Qy 3056 ATGAAAAAGCCCTGCTGATGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 3115
Db 3136 ATGAAAAAGCCCTGCTGATGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 3195
Qy 3116 GAGGAATGATGATTTAATGAGTGTGTTCACTATCTCCACTGAGCAGCAGCAGCAG 3175
Db 3196 GAGGAATGATGATTTAATGAGTGTGTTCACTATCTCCACTGAGCAGCAGCAGCAG 3255
Qy 3176 GTGCTCAGGGGCTGAGCTGAGTGAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 3235
Db 3256 GTGCTCAGGGGCTGAGCTGAGTGAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 3315
Qy 3236 AGTGGCTGGGAG 3295
Db 3316 AGTGGCTGGGAG 3375
Qy 3296 GGTTCAGTCTAATGATGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 3355
Db 3376 GGAACAGTGTCTTATGATGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 3435
Qy 3356 CACCTGGGATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 3415
Db 3436 ---CTGGGATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 3495
Qy 3416 GCTATGAG 3475
Db 3493 GCTATGAG 3555
Qy 3476 GGCACATGACATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 3535
Db 3553 GGCACATGACATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 3615
Qy 3536 AAAGGAAG 3595
Db 3613 AAAGGAAG 3675
Qy 3596 AGACAGCTCATTTTGTGATGATGATGATGATGATGATGATGATGATGATGATGAT 3655
Db 3673 AGACAGCTCATTTTGTGATGATGATGATGATGATGATGATGATGATGATGATGAT 3735
Qy 3656 AAGGTGTCAG 3715
Db 3733 AAGGTGTCAG 3795
Qy 3716 CACGCTGTCAG 3775
Db 3793 CACGCTGTCAG 3855

Qy 3776 AAGAGCATGCGACACATCAACACAGCTGCGCCGACAGAAAGCAGATTTTTCATGCTC 3835
Db 3853 AAGAGCATGCGACACACATCAACACAGCTGCGCCGACAGAAAGCAGATTTTTCATGCTC 3909
Qy 3836 ACTGTCAG 3895
Db 3910 ---GTTGAG 3965
Qy 3896 TTAATATTTGTTTAAACATGAGATTTAATCAAAAGTTAAAGGTGA 3942
Db 3966 CTAATATTTGTTTAAACATGAGATTTAATCAAAAGTTAAAGGTGA 4012

RESULT 12
US-09-880-107-2299
; Sequence 2299, Application US/09880107
; Patent No. US20020142981A1
; GENERAL INFORMATION:
; APPLICANT: Horne, Darci T.
; APPLICANT: Vockley, Joseph G.
; APPLICANT: Scherf, Uwe
; APPLICANT: Gene Logic, Inc.
; TITLE OF INVENTION: Gene Expression Profiles in Liver Cancer
; FILE REFERENCE: 44921-5028-WO
; CURRENT APPLICATION NUMBER: US/09/880,107
; PRIOR APPLICATION NUMBER: US 60/211,379
; PRIOR FILING DATE: 2000-06-14
; PRIOR APPLICATION NUMBER: US 60/237,054
; PRIOR FILING DATE: 2000-10-02
; NUMBER OF SEQ ID NOS: 3950
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 2299
; LENGTH: 3924
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; OTHER INFORMATION: Genbank Accession No. US20020142981A1 M23234
US-09-880-107-2299

Query Match 52.3%; Score 2240; DB 10; Length 3924;
Best Local Similarity 74.8%; Pred. No. 0;
Matches 2887; Conservative 0; Mismatches 945; Indels 30; Gaps 5;
Qy 50 GAGAGACTTGTGAGAAATGGCCAAAGTAAAGTAAAGTAAAGTAAAGTAAAGTAAAGTAA 109
Db 81 GCGAGAGGAGACTTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAG 140
Qy 110 ---CCAACTGTGAGCAGCTTTCGCAATGTTTGGCTATTCAAATTTGGCTGATAGGTTGAT 166
Db 141 GTGAAATATGATGAGTATTAATCAATGTTTGGCTATTCGATTCGATGAGATTAATTTGTTT 200
Qy 167 ATGTTGGTGGGAGCATGAGCTGATCATCATGAGAGTGCAGTCCCTCTCATGATGCTG 226
Db 201 ATGTCGTGGGTACATCATATGAGCTGATGAGTGCAGTCCCTCTCATGATGATGATGAT 260
Qy 227 GTTTTGAAGAGTGAAGATGAGTTCGAAATGCAAGATTTGCAAGAAACAAACTTTT 286
Db 261 GTATTTGAGAGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 314
Qy 287 CCAATTAATTAATGAAATATTAAGAAATATTAAGAAATATTAAGAAATATTAAGAAAT 346
Db 315 CCAATTAATTAATGAAATATTAAGAAATATTAAGAAATATTAAGAAATATTAAGAAAT 362
Qy 347 GAGGAATGAGAGAGATGAGTATTAATTAAGAGTGGATGAGTGGTGGTGGTGGTGGTGGT 406
Db 363 GAGGAATGAGAGAGATGAGTATTAATTAAGAGTGGATGAGTGGTGGTGGTGGTGGTGGTGGT 422
Qy 407 GCTTACATCAGGTTTCAATCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 466
Db 423 GCTTATATCAAGTTTCAATCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 482

QY	467	AAACAATTTTTTCAATGCTATCATATGCACAGAGAGATTGGCTGGTTGAGCTGATGACGTT	526
Db	483	CAGAAAGTTTTCATGCTATCTTACAGACAGAAATAGATGGTTTGACATCAATAGACACC	542
QY	527	GGGAGCTTAAACCCGGGCTCACAGACAGATGCTCCAAAATCAATGAAGCAATTGGCGAC	586
Db	543	ACTGACATCAATACGGCGCTTACAGAGATGACATCTCCAAAATCAGTAGAAGCAATTGGAC	602
QY	587	AAAAATGGAAATGTTCTTTCACTCAATAGCAACAATTTTTACCGGTTTTATATAGGGGTTT	646
Db	603	AAGGTGGAAATGTTCTTCAAGCAGTAGACACGTTTTTTGCAGGATTCATAGGGGAATTC	662
QY	647	ACAGGTGTTGGAAGCTACACCCCTGTGATTTTGGCCATCAGCCCTGTTCTTGACCTTTCA	706
Db	663	ATCAGAGAGATGGAAGCTACACCCCTGTGATTAATGGCCATCAGCCCTATTTAGGACTCTCT	722
QY	707	GCCGCATCTGGGCAAAAGATACATATCTTCATTTACTGATTAAGAACTCTTGGCCTATGCA	766
Db	723	GCACCCGTTTGGCCAAAGATACTCTGGCATTTAGTGCAAAAGAACTGCTGTTATGCA	782
QY	767	AAAGCTGGACAGTACCTGAGAAGTCTTTAGACAAATCAGAACTGATGATGGCTTTGGA	826
Db	783	AAAGCAGGCCCGCTGGCAGAAAGAGCTCTGGGGGCCATCAGAGCTGATAGCTTTGGG	842
QY	827	GGAAAAAGAAAAGAACTTTAAAGATCAACAACAAATTTAGAAGAGCTAAAGCAATTGGG	886
Db	843	GGCCAGACAAAGAGCTGGAAAGATATCAGAAACATTTAAACAAATCCCAAAGAGATTGA	902
QY	887	ATTAAGAAAGCTATCACGGCCAAACATTTCTATTGGTCCGCTTCTGTTATGATCTATGCA	946
Db	903	ATTAAAAAGCTATTTCAGCAAAACATTTCCATGGGTATTCCCTTCTGTTAATATATGCA	962
QY	947	TCATATGCTCTGGCTTTCTGGTATGGAGCTCCTTGGTCTCTCCAGTAAATATTTCTATT	1006
Db	963	TCATATGCACTGGCTCTCTGGTATGGATTCACACTCTATGCTATCAAAAGAAATATCTATT	1022
QY	1007	GGACAAGTACTACTGCTCTCTCTTTCTGCTATTAATTTGGGGCTTTAGTATTTGACAGCA	1066
Db	1023	GGAAATGCAATACAGTTTTTTTTTTTCAATCCTAATTTGGAGCTTCAAGTTGGCCAGGCT	1082
QY	1067	TCCCCAAGCTTTGAAGCATTTTGGCAAGCAGAGAGAGACAGCTTATGAATCTTCAAGTA	1126
Db	1083	GCCCACTGATTTGATGCTTTTGGCAATGCAAGAGGAGCAGCATATGTATCTTTGATATT	1142
QY	1127	ATTGACAATTAACCCAGAGATTTGAGAGCTATTTCAGAAAGTGGACATTAACAGATTAATT	1186
Db	1143	ATTGATTAATATCCTAAATATGACAGTTTTTTCAGAGAGGAGACAAACACAGACAGATC	1202
QY	1187	AAAGGAAATTTGGAATTCAAAATAATGTTCACATTAAGTACCCCTTCGCAAAAAGATTAG	1246
Db	1203	AAAGGAAATTTGGAATTCATATGATGTTCACTTTTCTTACCCCTTCGAGCTTAACGTAAG	1262
QY	1247	ATCTTAAGGCTCTCAACTGTAAGGTTCAAGAGTTGGGACAGACAGTGGGCTGTTGGGAC	1306
Db	1263	ATCTTAAGGAGGCTCTCAACTGTAAGGTTGGAGAGTGGGACAGAGTGGGCTGTTGGAGT	1322
QY	1307	ACTGCTCCGGAGAGACACAGACCGTGGAGCTGATGCGAGAGGCTCTATGACCCACAGAT	1366
Db	1323	ACTGCTCTGGGAAGACACAAAGCTCCAGCTGATACAGAGGCTCTATGACCCCTGATAG	1382
QY	1367	GGCATGCTCTGATTTGATGAGACAGCAATTAGACCAATAATGTAAAGCACTTTGGGAA	1426
Db	1383	GGCACAATTAACATTTGATGGCAGAGATTTTAGAATCTTAATGTAACTATCTGAGGAA	1442
QY	1427	ATTACTGGTGGGAGAGTCAAGAGCCTGTGTTTGGCCACACAGATAGCTGAAAAATT	1486
Db	1443	ATCATTTGGTGGTGGAGTCAAGAGCCGGTGTGTTTTCACACAAATTTGCTGAAATATT	1502
QY	1487	CGCATGCGCGAATAATGTCAACATGATGATTTGAAAGAGCTGTTAAGAAAGCCAT	1546
Db	1503	TGTTATGGCGTGGAAATGTAAACCATGATGATTAAGAAAGCTGTCAAAAGGGGCAAC	1562
QY	1547	GCCATATTTTTCATGAAGCTACCTAATAAATTTGACACTGTGTTGGAGAGAGAGG	1606

Db	1563	GCCTATGAGTATTATCATGAATTTACACAGAAATTTTACACACCTGGTTGGAGGAGGG	16
QY	1607	GGCCAGCTCAGTGGTGGACACAAACAGAGAAATGCCATTGCTGTGGCCCTGTTGGCAAC	16
Db	1623	GGCCAGCTCAGTGGTGGACACAAACAGAGAAATGCCATTGCTGTGGCCCTGTTGGCAAC	16
QY	1667	CCCAAGATTTCTTGTCTGGATGAGGCAACGCTCAGCTGTGGACCTGAAGTAAACACATG	17
Db	1683	CCCAAGATTTCTTGTCTGGATGAGGCAACGCTCAGCTGTGGACCTGAAGTAAACACATG	17
QY	1727	GTTCAGGTGGCCCTGGATAGGCCAGAAAGGCCCGGACTACCATTTGTATAGCTCATGCT	17
Db	1743	GTACAGGAGAGCTGTGGATAGGCCAGAAAGGCCCGGACCACCATTTGTATAGCACACCGA	18
QY	1787	TTGCTACAGTTCCTAATGGCGSAGTGCATGCTGGTGTGATGATGAGTCAATTTGGAG	18
Db	1803	CTGTCTACAGTTCCTAATGGCGSAGTGCATGCTGGTGTGATGATGAGTCAATTTGGAG	18
QY	1847	AAAGGAATCATGTGATGACTGCATGAAAGAGAAAGGGCATTTAGTTCAAACTGTGCACATG	19
Db	1863	CAAGGAAGCCACAGCAAGTGCATGAAAGAGAAAGGGGTGCTACTTCAAACTGTGCACATG	19
QY	1907	CAGACAGAGGAAATGAAATGATGTTAGAAATGCCATCGTGAATCCAAAAGTAAAGT	19
Db	1923	CAGACATCAGGAAAGCCAGATTCAGTGAAGA- ----ATTGAACTAAATGATGAAGAAG	19
QY	1967	GATGCTTGGAAATGTCCTCCAAAGATTCAGGGTCCAGCTTAATPAAAAAGAAATCACT	20
Db	1977	GCTCCCACTAGAAATGGCCCCAAATGGCTGGAATCTGCCCTTTTAGCATTTTACTCAG	20
QY	2027	CGCAGAGTATACATGCACACACAGGCCACAGAAAGCTTGGTACAAAAGAGCACTTG	20
Db	2037	AAAAAAGCTTAAAAATTCACAAATGTGTCAAGAAAGCCTTGATGTGAAACCGAGTGCCT	20
QY	2087	AATGAGATTTACCTCCAGTTCCTCTGGAGAGATTCGAGAGTGAAGTCACTCACTGAATG	21
Db	2097	GAGGCAAAATGTGCCACACACTGTCTTTCTGAAGTCTCTGAAGTCAAGTAAAAAGAAATG	21
QY	2147	CGTATTTTGTGGTGTGATATTTTGTGCTATATTAACGAGAGCCTCGACACAGATTT	22
Db	2157	CCCTACTTTGTCTGTGGAAACAGTATGTCCATTCGCAATGGGGGCTTACGCCGATTT	22
QY	2207	TCATATATATTTTTCAGAGATTTATAGGAGTCTTTACCCGATGAGATCCTGAACAAA	22
Db	2217	TCAGTATATTTCTCAGAGATTCATACGATTTTGGACAGGCAATGA- --TGCAGTGAAG	22
QY	2267	CGACAGAGTATACATGTTTCTGTATTTCTTGTAGTCTTGGAAATTTATTTCTTTAT	23
Db	2274	CACGACAGATGCAACATATTCCTTTGATTTTCTTATTTTGGGAAATTTATTTCTTTTT	23
QY	2327	ACATTTTCTCCAGGCTTCACAATTGGCAAGCTGGGAGATTCCTCACTAAAGCGGCT	23
Db	2334	ACTTTCTCTTTCAGAGGTTTCACTTTTGGGAAAGCTGGCAGAGATCTCACAGAAAGCTG	23
QY	2387	CGATACATGTTTATAGATCATGCTGAGACAGAGATGACGCTTGTATATACCTTAA	24
Db	2394	CGGTCAATGGCTTTTAAAGCAATGCTTAAGACAGACATGAGCTGTGTTTATACATATA	24
QY	2447	AACCACTGAGACATTGACAAACAGGCTTGCCAATGATGCGGCTCAATTTAAAGGGCT	25
Db	2454	AACCACTGAGGACCTTTTCACAAAGACTTGCACAGATGTGTCGCCAATGCCAAGAGACC	25
QY	2507	AATGCTTCAGGCTTGCCTATTTACCCAAATATATAGCAAAATCTTTGGCAGGCAATAT	25
Db	2514	ACAGGAACAGGCTTGTATTTATTTGACAGAAATATGCTAACCTGTGTAACGTGTATATC	25
QY	2567	ATATCTTATCTATGTTGGCAATTTACACTTTTACTCTTAGCATTTGTACCATCAT	26
Db	2574	ATATCTTATCTATGTTGGCAATTTACACTTTTACTCTTAGCATTTGTACCATCAT	26
QY	2627	GCATAGAGAGATTTGTAATGAATATGTGTCTGCAAGCAAGCACTGAAGATTAAGAA	26

QY	180	CAATGCGTCCCAATCATCCATGAGAGCTGCACCTCCCTCTCATGATGGCTGGTTTTGGAAACA	235
Db	205	CCGCCATGGCAATAGCTCACGGAATCAGAGCTCTCCCTTATGATGATAGCTTTGGAGAAA	266
QY	240	TGACAGATAGCTTTGGCAATAGCAGGAATTTCAAGAAACAACCTTCCAGTTAATTA	295
Db	265	TGACAGATAACTTTGTAGAT-----AATGCTGGGAACTTTTCTTGGCAGT-----	310
QY	300	ATGAAGTATTAGCAACAAATACAAACATTTCTACACATCTGGAGGAGGAATGACA	359
Db	311	----GAATTTTCTTGTCTCAATGCTAAATCCAGGAAGAATTTCTGAGAGAGAAATVGTCTA	366
QY	360	CGATGCCATATTAATACAGTGGCATCGGTGCTGGCTGGTGGTCTGCTTACATCCAGG	419
Db	367	GATATGCACTACTATTTCTGGGACTAGGTGGTGGAGGTTCTTTGGCTCCATATTCACG	426
QY	420	TTTCATTTCTGGTGCCTGGCGACGAGAAAGACATCTCAAAATTAAGAAACATTTTTC	479
Db	427	TCTCCTTCTGGACTTTGGCAGCTGGCGCCGAATTAAGAAATACAGCAAAAATTTTTC	486
QY	480	ATGCTATCATCCGACAGAGAGATTTGGCTGTTTGACGTGCATGACGTTGGGAGCTTAACA	539
Db	487	ACGCAATCTTTCACAAAGAAATGGCTGGTTTGATATCAAGGCAACACGCACTCAACA	546
QY	540	CCCGGCTCACAGACGATGCTCCAAAATCAATGAAGGAATTTGGCGACAAATTTGAATGT	599
Db	547	CGCGCTGCACAGATGCATCTCTCAAAATACGTGAAGGAATTTGGTATGCAAGGTGGAATGT	606
QY	600	TCTTTCACTCAATAGCAACATTTTTCACCGGTTTATAGTGGGGTTTACACGTGTGTGA	659
Db	607	TCTTTCAAGCAATAGGCACGTTTTTTGGAGATTCATAGTGGGTTCAACAGGCTGGA	666
QY	660	AGCTAACCCCTGTGANTTTTGGGCATCAGCCCTGTTCTTTGGACTTCAGCCCATCTGGG	719
Db	667	AACTCACTCTCGTGAATCAGGCAACCGGCATCTGGGGCTCTCTACAGCCGTTGGG	726
QY	720	CAAAATACTATCTTCAATTTACTGATTAAGAAACCTTGGGCTATGCAAAAGCTGACAGAG	779
Db	727	CAAAATACTCTCAACATTTCACTAGTACAAAGAACTAGCTGCTATGTGAAAAGCAGGTGGCG	786
QY	780	TAGCTGAAGAACTTTAGCAGCAATCAGAACTGTGATTTGGCTTTGGAGGACCAAAAGAAG	839
Db	787	TGGCGGAAGAGCTCTGGGAGCCATTCAGGACCGGTGATAGCTTTCCGGGGCCGAAACAAG	846
QY	840	AACTTGAAGAGTACAAACAAATTTTAAABAACTTAAGGAATTTGGATTAAGAAAGCTA	899
Db	847	AGCTGAAGAGTATCAGAGACATTTTAAAAATCCAAAAAGATTTGAATTTAAAAAGCTA	906
QY	900	TGACGGCCACATTTCTAATTTGGCCGCTTTCTTATGATCATGATCATATCTCTGG	959
Db	907	TCTCGGCCAATCTCCATGAGGACCTTGCTTTTGTTAATATATCATCTATCATCTGG	966
QY	960	CTTTTGGTATGGGACCTCTTGGTCTCTCCAGTGAATATTCTAATTTGGACAACTACTCA	1019
Db	967	CCTTTGGTATGGATTCACATCTGGTTATTCAAAAGAAATATTCAAATTTGGAATGCCATGA	1026
QY	1020	CTGTCTCTCTTCTGATTTAATTTGGGGCTTTAATAGTAATGGACAGGCAATCCCAAGCATTG	1079
Db	1027	CAGTGTCTTCTCTCAATCTCATTTGGGGCCTTCAAGTGTGGGAGGAGTGGCCCCCTGATTTG	1086
QY	1080	AAGCATTTGCAAGCAGACAGAGACGCTTATGAATCTTCAAGATATTAATGACATTAAG	1139
Db	1087	ATGCTTTCCCAATGCTAGAGGAGCGACCTATGATCTTTGACATTAATTAATTAATATC	1146
QY	1140	CAACATTTGACACCTTTTGAAGAGTGGACATNAACCCAGATATATTAAAGGAAATTTGG	1199
Db	1147	CTAATAATGACATTTTTCAGAGAGAGGACCAACGCCAGACAGCATCAAGGAAATTTGG	1206
QY	1200	AATTCAAAAATGTTCACTTCACTTACCTTCCGTAACAAAGATTTAAGATTTAAAGGTC	1259
Db	1207	AGTTCAAGTACGCTTCACTTTTCTCAACCATCTCGGGGCTAATATCAAGATCTTGAAGGGCC	1266
QY	1260	TCAACCTGGAAGGTTCAAGATGGGCAAGACATGGCGCTGTGGTTGGGAACAGTGGCTGGGGA	1319

Db	1267	TCACCTGAAAGTGAAGCGGCGAGAGCGGAGCCCTGGTTGGCAACATGGCTGTGGGA	13
Qy	1320	AGGACGACGCGGTGACGTGATGCAGAGGCTCTATGACCCCAAGATGGCATGGTCTGTGA	13
Db	1327	AAAGCACAACCTGTCCAGCTGCTGAGAGGGCTTACAGCCCCACAGAGGGGTACGATTAGA	13
Qy	1380	TTGATGACAGAGACATTAGCACCTAAATCTAAAGGCATCTGGGGAAATTAAGTGGTGG	14
Db	1387	TCGATGGCGAGSACATCCGGAACTTTAACTCAGGGTGTCTAAAGGAAATTCATCGCGTGG	14
Qy	1440	TGAGTCAGAGACCCTGTGTGTTTGGCACCAAGATACCTGAAAAACATTCGCTATGGCGCG	14
Db	1447	TGAGTCAGAGACCCTGTGTGTCTACTGTCTACACGATGTCTGAAAATTCGGCTATGGCCGTG	15
Qy	1500	AAATGTACCATGATGATGATGAGAAAGCTGTAAAGAACCAATGCGTATGATTTTA	15
Db	1507	GGATGTATACATATGATGATGATTAATAAAGCTGTCAAAAGAGCTAATCCATATACATTCA	15
Qy	1560	TCATGAACCTACCTAATTAATTTGACACTGTGGTTGGAGAGAGAGGGGCCAGCGTGATG	16
Db	1567	TCATGAACCTGCCACAGAAATTTGACACCCCTGGTTGTGACAGAGGGGCCGACCTTGAGCG	16
Qy	1620	GTGACAGAAACAGAAATCCGCAATTCGCTGGGGCCCTGGTTGCCAACCCCAAGATTCCTC	16
Db	1627	GGGACAGAAACAGAGATTCGCAATTCGCTGCGCTTGGTCCGCAACCCCAAGATTCCTCC	16
Qy	1680	TGCTGATGAGGCAAGCTGACGTCTTGACACTGAAAGTGAAGCAAGTGGTTCAGGTGGCCC	17
Db	1687	TGCTGAGCAGGCGCAGCTGACCTTGTGACACAGAAAGGGAAGCTGAGGTGCAAGCCGCTC	17
Qy	1740	TGATTAAGGCAAGAAAGCCGCGACTACCATTTGATATGCTATCGTTTGTCTACAGTTTC	17
Db	1747	TGATTAAGGCAAGAAAGCCGCGACACCATCTGATATGCTATCGTTTGTCTACAGTTTC	18
Qy	1800	GTAATCCGATGTCATTCGTGTTTGTATGATGAGATCATTTGTGGAAAGAAATCATG	18
Db	1807	GGATGCAATGTCTCATTCGTGTTTGTATGATGAGATCGTCATTCGTGGCAAGAAAGCA	18
Qy	1860	ATGACACTGTAAGAAAGAAAGGCAATTTACTTCAAACTGTCACAATGCAACAGCAAGGAA	19
Db	1867	GTAAGCTGATTAAGAAAGAAAGGATCTACCTCAACCTGTTAACATGTGACACATCAGAA	19
Qy	1920	ATGAAATGATTAATAAATGCCACTGCTGATCCAAAGTGAAAGTGAATGCCCTGGAAA	19
Db	1927	GCCAGATCTGTCAAGAAATTTGAAAGTGGACTGATGATGAAGAAAGCGCTGGAGGTG	19
Qy	1980	TGCTCTCAAAAGATTCAGGGTCTCAGTTTAATTAATAAAGAAATCACTCCAGAGATATC	20
Db	1987	TGGCCCAAAATGGCTGGAAACACGCAATTTTGAAGATTTCAAGAAAGAAATCTGAAA	20
Qy	2040	ATGCACCAACAGGCCAGACAGAAAGCTTGATCAAAAAGSACTTGATGAGATGTAC	20
Db	2047	GTTTCAAGGCGCCATCAAAATATGAGCTGATGTGAACCAATGATGAAACGTGC	21
Qy	2100	CTCCAGTTTCCTCTTGAGAGATTTGAAAGTGTGAAGTGAACCACTGAATGGCTTATTTTGTGG	21
Db	2107	CACAGATCTCTTTCTTAAGTCTTAAAGATGATTAATAAACAAGTGGCCCTACTTGTGG	21
Qy	2160	TTGATTAATTTTGTCTAATTAATAAGAGAGGCTCGAACACACTTTTTCATTAATATTTT	22
Db	2167	TGGGAGCACTCTGGCAATGGCCAGACGGGGCCCTCCAGCGGCAATTCCTCAATCCCTGT	22
Qy	2220	CAAGATTAATTAAGSACTTTACCCGAGATGAGATCTGAAAACAACAGACAGATAGTA	22
Db	2227	CAGAGATGATGATTAATTTTGGCC---CTGGGAGTGACACAGTAAAGCAACAGAGGTGA	22
Qy	2280	ACATGTTTCTGTATTTGTTTCTAGTCCCTGGAAATTAATTCCTTAATTAACATTTTCTCC	23
Db	2284	ACATGTTTCTGTGTTCTTCTTGGGCTTAAGAGTCCACTCTCTTTCATCTTCTCTCTC	23
Qy	2340	AGGCTTCACTTTGGCAAGCTGGGAGATCTTCATTAAGCGGCTTCGATACATGGTTT	23

Db 2344 AGGGTTTCACATTCGGGAAAAGCTGGGAGATCCTCACACAAAGGCTCCGGTCCATGGGCT 2403
Qy 2400 TCAGATCCATGTCGAGACAGATGTCAGCTGGTTGATGATGCCCTAAACACACACTGGAG 2459
Db 2404 TCAAGCAATGCTAAGACAGATGATGCTGGTTGATGATGATGATGATGATGATGATGATG 2463
Qy 2460 CATTGACACACAGGCTTCCATGATGCGGCTCAATTAAGAGGGCTATAGGTTCCAGGCT 2519
Db 2464 CCTCTCTACAGACTGCGCACAGAGCTGCGAGGCTCCAGAGGCTCCAGAGGACAGGCT 2523
Qy 2520 TTGGTGCATTAACCAATATATGCAATCTTGGGACAGGCTATTAATATATCTTAATCT 2579
Db 2524 TGGCTTAATGTCACACAGACAGGCTTGGAGAGGCTATTAATATATCTTAATCT 2583
Qy 2580 ATGGTTGGCAATTAACACTTTTACTCTTACCAATTTGACCAATTCGCAATTCGCAATTC 2639
Db 2584 ACGGTTGGCACTGACACTTCTGCTTATCAGTTGTTTCATTCATTTGCTGTAGCGGGA 2643
Qy 2640 TTGTTGAATGAATAATGTTGTCGACAGCACTGAAGAGTAAGAAAGAGCTGAAGAG 2699
Db 2644 TTGTTGAATGAATAATGTTGTCGACAGCACTGAAGAGTAAGAAAGAGCTGAAGAG 2703
Qy 2700 CTGGAGATTTGCTACAGAGCCATTCGAAAATCTCCGACTGTTGTTCTTCTGACTCGG 2759
Db 2704 CTGGAGATTTGCTACAGAGCCATTCGAAAATCTCCGACTGTTGTTCTTCTGACTCG 2763
Qy 2760 ACAGAAATTTGAATACATGATGATGACAGAGATTTGCAAGTACATGACAACTCTTTGA 2819
Db 2764 ACAGAAATTTGAATACATGATGATGACAGAGATTTGCAAGTACATGACAACTCTTTGA 2823
Qy 2820 GGAAGACACACATCTCTGGGGCTCTCATTTCTATCACCAGGCAATGATGATTTCT 2879
Db 2824 GGAAGCTCACATCTGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 2883
Qy 2880 ATGCTGGCTGTTCCGGTGGTGGCTTGGCTTGGCTTGGCTTGGCTTGGCTTGGCTTGG 2939
Db 2884 ATGCTGGCTGTTCCGGTGGTGGCTTGGCTTGGCTTGGCTTGGCTTGGCTTGGCTTGG 2943
Qy 2940 ATGCTGGCTGTTCCGGTGGTGGCTTGGCTTGGCTTGGCTTGGCTTGGCTTGGCTTGG 2999
Db 2944 ATGCTGGCTGTTCCGGTGGTGGCTTGGCTTGGCTTGGCTTGGCTTGGCTTGGCTTGG 3003
Qy 3000 CATTGCTCTGCTATGCAAGCAAGGCTATGCAAGGCTATGCAAGGCTATGCAAGGCT 3059
Db 3004 CATTGCTCTGCTATGCAAGCAAGGCTATGCAAGGCTATGCAAGGCTATGCAAGGCT 3063
Qy 3060 AAAAAAGCCCTGATGATGACAGCTACAGCCCTCAGGCTCAGCCCAATACCTGGAAG 3119
Db 3064 AAAAAAGCCCTGATGATGACAGCTACAGCCCTCAGGCTCAGCCCAATACCTGGAAG 3123
Qy 3120 GAAATGATGATTAATGAGGTGCTGTTCACTATCCACTGACACAGATCCCGCTGC 3179
Db 3124 GAAATGATGATTAATGAGGTGCTGTTCACTATCCACTGACACAGATCCCGCTGC 3183
Qy 3180 TCCAGGGCTGAGGCTGAGGTGAGAGGCGCAGAGCTGGCCCTGTAAGTACAGTG 3239
Db 3184 TCCAGGGCTGAGGCTGAGGTGAGAGGCGCAGAGCTGGCCCTGTAAGTACAGTG 3243
Qy 3240 GCTGTGGAGAGACAGAGTGTTCAGTCTCCTAGAGGCTTCTAGAGCCCTTGGCTGTT 3299
Db 3244 GCTGTGGAGAGACAGAGTGTTCAGTCTCCTAGAGGCTTCTAGAGCCCTTGGCTGTT 3303
Qy 3300 CAGTGTGATGATGAGGCAAGATTAAGACACTGATGTCAGTGGCTCCAGACACC 3359
Db 3304 CAGTGTGATGATGAGGCAAGATTAAGACACTGATGTCAGTGGCTCCAGACACC 3363
Qy 3360 TGGGATGCTGTCAGAGAGCCATCTGTTGATCTGACATGTCGCGAATAATGCTG 3419
Db 3364 TGGGATGCTGTCAGAGAGCCATCTGTTGATCTGACATGTCGCGAATAATGCTG 3423
Qy 3420 ATGAGACAAAGCGGGGCTGATACATGAGATTAAGATTAAGAGGACCAAGAGGCA 3479
Db 3424 ATGAGACAAAGCGGGGCTGATACATGAGATTAAGATTAAGAGGACCAAGAGGCA 3483

Qy 3480 ACATACACACATTCATGAGACATCTCCTGAGAAATTAACACACAGATGAGCAAG 3539
Db 3484 ACATACACACATTCATGAGACATCTCCTGAGAAATTAACACACAGATGAGCAAG 3543
Qy 3540 GAAACACATCTCTGGTGGCGCAAGAACAGCGATTTGCCATTCGCGCTTGTAGAC 3599
Db 3544 GAAACACATCTCTGGTGGCGCAAGAACAGCGATTTGCCATTCGCGCGCATCATCAC 3603
Qy 3600 AGCTATATATTTGCTTTTGGATGAAGTACATCAGCTCTGATACAGAAAGTGAAGAG 3659
Db 3604 AGCTATATATTTGCTTTTGGATGAAGTACATCAGCTCTGATACAGAAAGTGAAGAG 3663
Qy 3660 TTGTCAGAAAGCGCTGGAGAAAGCCAGAGAGGCGCAGCTGCAATTTGATGCCACC 3719
Db 3664 TTGTCAGAAAGCGCTGGAGAAAGCCAGAGAGGCGCAGCTGCAATTTGATGCCACC 3723
Qy 3720 GCTTGTCCACCATTCAGAAATGATGATTAATGATGATGATGATGATGATGATGATG 3779
Db 3724 GCTTGTCCACCATTCAGAAATGATGATTAATGATGATGATGATGATGATGATGATG 3783
Qy 3780 AGCATGCAACATCAACAGTGTGCTGGCCAGAAAGCATCTATTTTCCATGCTCAGTG 3839
Db 3784 AGCATGCAACATCAACAGTGTGCTGGCCAGAAAGCATCTATTTTCCATGCTCAGTG 3843
Qy 3840 TCCAGGCTGAGAAAGCGCTGATGAACTGTCGCTGATGATGATGATGATGATGATG 3899
Db 3844 TCCAGGCTGAGAAAGCGCTGATGAACTGTCGCTGATGATGATGATGATGATGATG 3903
Qy 3900 TATTTGCTG 3908
Db 3904 TATTTGCTG 3912

RESULT 14
US-09-873-409-14
; Sequence 14, Application US/09873409
; Patent No. US20020037522A1
; GENERAL INFORMATION:
; APPLICANT: Frank, Markus
; APPLICANT: Sayegh, Mohamed
; TITLE OF INVENTION: A Gene Encoding a Multidrug Resistance Human P-Glycoprotein
; FILE REFERENCE: 81994/268611
; CURRENT APPLICATION NUMBER: US/09/873, 409
; CURRENT FILING DATE: 2001-06-05
; SOFTWARE: PatentIn version 3.0
; SEQ ID NO 14
; LENGTH: 3621
; TYPE: DNA
; ORGANISM: Homo sapiens
US-09-873-409-14

Query Match 30.9%; Score 1323.8; DB 10; Length 3621;
Best Local Similarity 62.6%; Pred. No. 2.8e-293; Indels 51; Gaps 6;
Matches 2189; Conservative

Qy 418 GATTTCATTCGCTGCTGCGACAGAGACAGATATCAAAATTAAGAAACAAATTTT 477
Db 162 CAGGTTCACCTGATTAATGTTGATAGTGTGCTGCTGATTTTGGTTACATACA 221
Qy 222 GATTTCCCTTGATTAATGTTGATAGTGTGCTGCTGATTTTGGTTACATACA 281
Db 478 TCATGCTATCAATGAGACAGAGATTTGCTGCTGATGATGATGATGATGATGATG 537
Qy 282 TCATGCTATTTTGGACAGACATGCTGCTGATGATGATGATGATGATGATGATG 341
Db 538 CACCGGCTCACAGACATGCTTCCAAATCAATGAAGATTTGGGACAAATTTGAAAT 597
Qy 342 CACTGCAATGACAGACATTTG--ACAAATCAGATGATGATGATGATGATGATGATG 398

QY	1672	GATTCTTCTCCTGGATGAGCAACGTCAGCTCTGACACTGAAAGGAAAGCAGTGGTTC	1731
Db	1479	GATTCTGATTTTATGATGAGGCTACGTCGCCCTGGATTCAGAAACCAATGCACTGCTTCA	1538
QY	1732	GGTGGCCCTGGATAGAGCCAGAAAAGCCGACCTACCATTTGTGATAGCTCATCTGTTGTC	1791
Db	1539	AGCTGCACCTGGAAAGGCGAGCAAGGTGGGACTCAATATCGGTGAGCACACCCGACTTC	1598
QY	1792	TACAGTTCTGTAATGCCAGATGCTATGCTGGTTTATGATGAGAGTCATTTGGGAAAG	1851
Db	1599	TACTATTCGAAGTCAGATTTTGATTTGTGACCCCTAAAGATGGAATGCTGGGGGAAAG	1658
QY	1852	AAATCATGTATGAACTCATATAAAGAGAGGCGCATTTACTTCAAACTGTGCACATGCAAG	1911
Db	1659	AGCACATGCTGAACATTAATGGCAAAACGAGGCTCATATATATTCACCTTGTGATGTCAGGA	1718
QY	1912	AAGAGGAATGAAATGTAGATTGAAATAATGCCACTGCTGAAATCCAAAMGTGAAGTATGC	1971
Db	1719	TATTAATAAAGCTGATGA-ACAGATGAGAGTCAATGACATATTTCTGTAAGAAAGACCA	1777
QY	1972	CTTGGAAATGCTCCAAAAGATTCAGGGTCCAGTTTATATAAAGAGATCAACTGCGAC	2031
Db	1778	ACTCACTTCTCTGCACTGCTGGAAG-----CATCAAGTCAGACTTCATTTAGCA	1829
QY	2032	GAGTATACATGACACACAGGCGCAAGACAGAAAGCTGTATCAAAAGAGACTTGAATGA	2091
Db	1830	GGCTGAGAAATCCACCCAACTMAAGATTAAGCTT-----	1866
QY	2092	GAATTAACCTCCAGTTCCTCTTGAGAGATTGTGAAGCTGACATCAATGAATGGCTTGA	2151
Db	1867	-----CCTGAATCTCTCATATAAATTTAAAGTTAAACAGCTGAAATGGCTTT	1919
QY	2152	TTTGTGGTTGGTATATTTTGTGCTATTAATAAGGAGGCGCTCAACACAGATTTCAAT	2211
Db	1920	TGTGGTCTGGGACATTTGGCTTCTGTCTTAATGGAACGTGTCATCCAGTATTTTCCAT	1979
QY	2212	AAATATTTTCAAGATTATAGGATCTTTACCCGAGATGAGGATCCGAAAGCAAAACGACA	2271
Db	1980	CATCTTTGGCAAAATTAATTAACCATGTTT---GGAATATATGATTAACACATTTAAGCA	2036
QY	2272	GAATAGTAACATGTTTCTGATTTGTTTCTAGTCCCTGGAATTTTCTTTTATTAACAT	2331
Db	2037	TGATGCAAAATTTATTTCCATGATATTTCTCATTTTGGGTGATTTGGTTGTGCATTA	2096
QY	2332	TTTCTCCAGGCGCTTCACATTTGGCAAACTGGGAGATCCTCACTAAGCGGCTCGTA	2391
Db	2097	TTTTCATGCAAGGATTTATTTTACGGCGAGCAGGAGGAATTTACGATGAGATTTAAGCA	2156
QY	2392	CATGAGTTTCAGATTCACCTGACAGACAGATGTCAGCTGTGTTGATGACCTTAACAC	2451
Db	2157	CTTGGCGCTTCAAAAGCCATTTTATATCAGGATTAATCCCTGGTTGTGTAAGAAAGAAACAG	2216
QY	2452	CACGTGAGCATTTGACAAACCGAGGCTGGCAATGATCGGCTCAAGTATAAGGGGCTATAGG	2511
Db	2217	CACAGAGGCGCTTGACAAACAATATATAGCCATGATTAACACAAATTCAGAGAGCAACGG	2276
QY	2512	TTTCCAGGCTGTGTCATTAACCCAGAAATATATGCAATCTTGGACAGGCAATTATATATC	2571
Db	2277	TTTCCAGGATTTGGCGTCTTAACCAAAAGCAATTAAGGCACTTCAATTTATCATTTTC	2336
QY	2572	CTTAAATCTATGTTGGCAATTAACCTTTTACTCTTTAGCATTTGTACCCATTCATTCGAT	2631
Db	2337	CTTATATATATGATGGAGATGACATTTCTTATCTTGTAGATTTTCTCTCACTACTGCGGT	2396
QY	2632	ACCAGAGTGTGTGAATAAGAAATGTTGTCTGGAACAAGCACTGAAGATTAAGAAGAGCT	2691
Db	2397	GACAGGATTTGTTGAACCCGAGCAATGACGTGATTTTGGCCAAAGAAATTAAGCAAGACT	2456
QY	2692	AGAAAGAGCTGGGAAGATTTGCTACAGAGCAATCGAAAAATTCGGAACGTGTGTTCTTT	2751
Db	2457	TAAAGATCTCTGGAAGATAGCAACTGAAAGCTTTTGGAAATATACGATATATGCTGAT	2516
QY	2752	GATCGGAGCAAGATTTGAATACATGTATGCAACAAGATTTTGGCAAGTCCATTAAGAAA	2811

Db 2517 AACAGAGGAAAAAGCCTTCAGACAAATGTATAGAGAGTCTTCAGACTCAACACAGAAA 2576
QY 2812 CTCTTGAGAGAAAGACACATCTCGGGCTCATTTCTATACACCGCAATGATGTA 2871
Db 2577 TACCTGCAAAAGACACAGATATTTGGAAGCTGTTATGATTCAGCATGCTTATATATA 2636
QY 2872 TTTTTCATATGCGCTGCTTTCGGTTGTCGCTACTGTTGGGCAAAATGATTCATGAA 2931
Db 2637 TTTTGCTATGACAGAGGTTTGATTTGAGCTATTTAATTCAGCTGACGAAATGAC 2696
QY 2932 CTTTCAGAGTCTTTTGGTATTCAGACTATGCTTTGGTGCATGAGCAGTGGGCA 2991
Db 2697 CCCAGAGGGCAGTGTCAATAGTTTACTGCAATTCATATGAGCATGAGCATGCGAAA 2756
QY 2992 GGTGAGTTCAATTTCTCTCTGACTATGCAAAAGCCAAAGTATCAGAGCCCGCATCAT 3051
Db 2757 AACGCTGTTTGGCTCTGAAATATTCAAAGCCAAATCGGGGGCTGCGCATCTGTTTCG 2816
QY 3052 GATCATTTAAAAAAGCCCTCGATTTGACGTACAGCCCTCAGGCGCTCAAGCCAAATAC 3111
Db 2817 CTGTGTTGAAAAGAAACCAATATAGACAGCCGACGTCAAGAAAGGAAAAAGCAGACAC 2876
QY 3112 GTTGAAGAAATGTGACATTTAATGAGTCGTGTTCACTATCCACTGACAGACAAAT 3171
Db 2877 ATGTGAAGGAATTTAGAGTTTCGAAAGTCTCTTCTCTCATCATGATGCGCCAGATGT 2936
QY 3172 CCCCCGTCTCCAGGGGCTGAGGCTCGAGGTGAGAGAGGCGCAGCGCTGGCCCTGTAGG 3231
Db 2937 TTTTCATCTCCCTGGCTTATTCCTCATGATTTAGACGAGAAAGAACATGATTCATTTGGG 2996
QY 3232 TAGCAGTGGCTGTGGAGAGACAGTGTTCAGTCTCTAGAGCGCTTCTATAGCCCTT 3291
Db 2997 GAGCAGCGCTGTGGGAAAAAGCACTTCTGTCACTTCGTGAGAGACTTATATGACCCCT 3056
QY 3292 GCGTGTGATGCTATATGATGAGCAAAAGATTAAGCACTGATGCTCAGTGGCTCG 3351
Db 3057 GCAGAGCAAGTGTGTTGATGATGATGATGATGATGATGATGATGATGATGATGATG 3116
QY 3352 AGCAGACCTGGGCACTGCTCTCAGAGAGCCATCTGTTTGTAGCTGACAGATTCGCGAGA 3411
Db 3117 TTTCCCAATATGCAATGCTTCTCAAGAGCTGCTGCTTCACTGACATGCTGCTGAGAA 3176
QY 3412 CATTCCTATGAGACAAACAGCGGCTGTATCAGATGAGAGATTTAGCAGAGCCAA 3471
Db 3177 CATGCTTATGATGACAAAGCGGTGTGCTGATTTAGATGATGATCAAGAAAGCCGCAA 3236
QY 3472 GAGAGCAACATACACACTTCATGAGACACTCCCTGAGAAATACACACCAAGTATAG 3531
Db 3237 TGCAGCAAAATTCATTTCTTTATGAGGTCTCCCTGAGAAATACACACCAAGTATAG 3296
QY 3532 AGACAAAGAACCCAGCTCTCTGTGGCCAGAAAGAGCCATTTGCCATGCTCGGCTCT 3591
Db 3297 ACTAAAGAGACACAGCTTCTGTGGCCAGAAAGAAAGCTATGATGATGATGATGATG 3356
QY 3592 TGTAGACAGCTCATATTTTGTGTTTGTATGATGATGATGATGATGATGATGATGATG 3651
Db 3357 TCTCAAAAACCCAAATTTTATTTATTTGATGATGATGATGATGATGATGATGATGATG 3416
QY 3652 TGAAGAGTTGTCCAGAGAGCCCTGAGCAAAAGGCAAGAGCCGACCGCATGTTGTAT 3711
Db 3417 TGAGAAAGTGTGTCAGAGTGCCTTTGATTAAGCCAGAGAGGAGAGCATGCTATGTGT 3476
QY 3712 CGCCGACGCTGTGTCACATTCAGATGCAATGCAATTTAATAGTGTGTTTCAGATGGCAA 3771
Db 3477 CACTCAGAGGCTCTCTGTAATTCAGAAAGCAGATTTATGATGATGATGATGATGATG 3536
QY 3772 AGTCAAGAGGATGACACATCAACAGCTGCTGCCCCAGAAAGGCAATTTATTTTCAT 3831
Db 3537 GATTAAGGAAACAGAAATCATCAAGAGCTCTCTGAAATTCAGACATATATTTTAAGTT 3596
QY 3832 GGTGAGTGTCCAGGCTG 3848

Db 3597 AGTGAATGCACAGCTCAG 3613

RESULT 15
US-09-873-409-13
; Sequence 13: Application US/09873409
; Patent No. US20020037522A1
; GENERAL INFORMATION:
; APPLICANT: Frank, Markus
; APPLICANT: Sayegh, Mohamed
; TITLE OF INVENTION: A Gene Encoding a Multidrug Resistance Human P-Glycoprotein
; FILE REFERENCE: 81994/268611
; CURRENT APPLICATION NUMBER: US/09/873,409
; NUMBER OF SEQ ID NOS: 19
; SOFTWARE: PatentIn version 3.0
; SEQ ID NO 13
; LENGTH: 3702
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: NO. US20020037522A1e
; LOCATION: (723)..(723)
; OTHER INFORMATION: n at position 723 represents any nucleotide (A, T, C or G)
US-09-873-409-13

Query Match 28.8%; Score 1232.8; DB 10; Length 3702;
Best Local Similarity 61.2%; Pred. No. 1,9e-372;
Matches 2189; Conservative 0; Mismatches 1257; Indels 132; Gaps 7;

QY 358 CACGTATGCTATTTTATACAGTGGAGTCGTCGGGCTGCTGGGCTCTTACATCA 417
Db 162 CAGGTTGACCCGTATTTGTTGATATGATGATGATGATGATGATGATGATGATGATG 221
QY 418 GGTTCATCTGCTGCTGCTGAGCAGAGAGACAGATCTCAAAATTTGAAAACAAATTT 477
Db 222 GATTTCTGTTGATTTATTAATCACTGACAGCAGACAGCAAGAGATTTGAAAACGTTT 281
QY 478 TCATGCTATCATGACAGAGAGATTTGCTGTTGATGATGATGATGATGATGATGATG 537
Db 282 TCATGCTATGTTGACAGAGAGATGCTGTTGATGATGATGATGATGATGATGATGATG 341
QY 538 CACCGGCTCAGACAGAGATGCTCCAAATCATGAAAGATTTGGCAGCAAAATTTGAAT 597
Db 342 CACTGCATGACAGACATTTG---NCAAAATGATGATGATGATGATGATGATGATGATG 398
QY 598 GTTCTTCACTCAATAGCAACATTTTTCACCGGTTTATGAGGGGCTTTACAGCTGTTG 657
Db 399 GTTGTTCAAAACATGCTACTTCTTTCATGTTGCTGCTGCTGCTGCTGCTGCTGCTGCT 458
QY 658 GAAGCTAACCTGTTGATTTTGGCCATACGCCCTGTTCTTGAATTCAGCCGCATCTG 717
Db 459 GAAAGCTACCTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 518
QY 718 GGCAGAACTACTTCTTCAATTTACTATTAAGAACTTTGGCTTGTGCAAAAGCTGAGC 777
Db 519 TTTAGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 578
QY 778 AGTAGCTGAAGAGTCTTACAGCAATCTGAGAACTGATTTGCTTTGAGAGCAAAAGAA 837
Db 579 TGTGCGAAGAGAGTCTTGTTCATCAATCCGAAAGCATATGAGCTTTAGGCGCCAGAGAA 842
QY 838 AGAAC----- 842
Db 639 AGAAGCTCAAGGCTTCTTCTTTAATATATACAGATATGCTGTTTATTTTCCCA 698
QY 843 -----TTGAAGGTACACAAATAATTTAGAGAGACTTA 876
Db 699 GTGGCTACTAAGTTGTGTTCTGTTTNTTTGTAAGTATACAGAAATCTCAAAAGCTCAA 758
QY 877 AGGAATGGATTAAGAAAGCTATACGGCCACATTTCTATTTGGTGGCGCTTTCTATT 936

D	759	GGATTTTGGCAATAAAGAGACATAAGCTTCAAAAGTGTCTTGGTGTGTACTTCTT	818
Q	937	GATCATGACATCAATGCTGTGGCTTTCTGTATGGAGACCTCTGTCTCTCCAGTGA	996
D	819	TATGATGAGAACTATGAGACTTGTCTTTGGTATGGAACTCTCTGATCTTAATGAGA	878
Q	997	-----ATATTCATTTGGACAAAGTACTCACTGTCTTTCTGTATTAATGGGGCTT	105
D	879	ACCTGATATATACATGGGGGACCTGTCTGTCTTCTTACTGTATCATATGACATGA	938
Q	1051	TAGTATTTGGCAGCAGCCATCCCAAGCATTTAGACATTTGCAAGGACAGAGACGCTTA	1110
D	939	TTGCATTTGGAGCAGCAGCTCCCTCACTTCTTAACCTTCGCAATAGCCGAGAGCTGT	998
Q	1111	TGAAATCTTCAAGATTAATGACAAATTAACAGACATTTGACACTTTTGGAAAGTGA	1170
D	999	TCATATTTTCCAGGTATTGATTAAGAAACCCAGTATAGTAACTTTTCCACAGCTGGTA	1058
Q	1171	TAACACAGATTAATTAAGGAAATTTGGAATTCAAAAATGTCACTTCAGTTACCTTC	1230
D	1059	TAACCTGAAATCCATAGAAAGAACTGTGGAATTTAAAAATGTTCTTCAATTAATC	1118
Q	1231	TCGAAAAGATTTAAGATCTTAAAGGCTCAACCTGGAAGTTGACAGTGCAGAGCACT	1290
D	1119	AAGACCATCTATCAAGATTTGAAAGTGCTAACTCAGATTAAGTCTGGAGGACAGT	1178
Q	1291	GGCGCTGGTGGGACAGTGGCTGGGGGAAAGCAGCAGCCTGCAGCTGATGCAAGCT	1350
D	1179	GGCTTTGGTGGCTCAATAGCAGAGGGGAAAGATGAGTCCAGCTTCCGACAGGTT	1238
Q	1351	CTATGACCCACAGATGGCATGCTGTATTTATGACAGACAGCATTTGACCAATATG	1410
D	1239	ATATATCCGATGATGGCTTTATCATGTGGATGAGAAATGACATCAGACCTTAATGT	1298
Q	1411	AAGCATTTTGGGAAATTAAGTGTGGTGGAGTCAAGACCTGTGTGTTGCCACAC	1470
D	1299	GGCGATTTATGACACCATATTTGAGTGGTATGACAGACCTGTGTTGTCGGGACAC	1358
Q	1471	GATAGCTGAAACATTCGCTATGGCCCGACGAAATGTCAACCATGGATGATGAGAAAC	1530
D	1359	CATCAGTACCAATATCAAGATGAGACAGATGATGTACATGATGAAAGATGAGAGAC	1418
Q	1531	TGTTAAGGAGCCATGCTATGTTTTTATCATGAACATCACTAATTAATTTGACACTCT	1590
D	1419	AGCAAGGAGCAATGCTATGATTTTATCATGGAGTTTCCATTAATTAATTAATCAAT	1478
Q	1591	GGTTGAGAGAGAGGGGCCAGCTGAATGGTGGACAGAAACAGAAATGGCATTGCTG	1650
D	1479	GGTAGGGGAAAAAGGAGCTCAATATGATGGAGGGCAAAACAGAGATGCAATGTCTG	1538
Q	1651	GGCCCTGGTTCGCAACCCCAAGATTTTCTGCTGGATGAGGCAACGTACGCTGGACAC	1710
D	1539	TGCTTATGTTCCAAACCCCAAGATTTCTGATTTTAATGAGGCTACGCTGCCCTGATTC	1598
Q	1711	TGAAGTGAAGCAGTGGTTCAAGTGGCCCTGGATTAAGGCAAGAAAGCCGAGCTACAT	1770
D	1599	AGAAACCAAGTACGCTTCAAGTGCAGCTGAGAGAGGCAACAAAGGTCGACTACAT	1658
Q	1771	TGAGATGATCAGTCTGTTCTCTACAGTTCTGTATGCGGATGCTCATTTGCTGTTTGA	1830
D	1659	CGTGTAGACACCGACCTTTCTACATTTTCAAGTGCAGATTTGATTTGAGCCCTAAAG	1718
Q	1831	TGGAGTCATTTGAGAAAGAAATCATGATGAATCATGAAAGAGGCACTTTACTT	1890
D	1719	TGAGATGCTGGGGAGAAAGGAGCAGTCTGAATTAATGCAAAACAGAGCTATATATTA	1778
Q	1891	CAAAATTCACAAATGACGACACAAGGAATGAATTTGAGTTAGAAATGCCACTGTGTA	1950
D	1779	TTCACTTTGATGTCACGATATTAATAAAGCTGATGA-ACAGATGAGATCAATGACAT	1837
Q	1951	ATCCAAAGTGAAGTGTGCTTGGAAATGTTCCCAAAAGATTCAGGGTCCAGTTTAAT	2010
D	1838	ATTCTACTGAAAGAAAGCCACACTCTCTCTGCACTCTGTGAAG-----CAT	1889

QY	2011	AAAAAGAGATCAACTCGCAGGAGTATACATGCACACAGGCCAAGACAGAAAGCTTGG	2070
Db	1890	CAAGTCAGACTTCATTGACAGAGCTGAGGAATTCACCCAAATTAAGAGATTAAGCTT -	1947
QY	2071	TACAAAAGAGACTTGAAATGAGAAATGTACCTCCAGTTTCCTTGTGAGAGATTGAAAGCT	2130
Db	1948	-----CCTAAAGTCTCTCATTAATAAATTTTAAAGTT	1979
QY	2131	GAACTCAGCTGATATGGCTTATTTTGTGGTGTGTAATTTTGTGATATATAAACGAGG	2190
Db	1980	AAACAGACCTGATAGGCCCTTTTGTGTCTGTGGGACATAGGCTCTCTGTCTTAAATGGAAC	2039
QY	2191	CCGTCACACAGCATTTTCAATAATATTTTCAAGATATTAAGGATCTTTACCCAGATGA	2250
Db	2040	TGTCATCCAGATATTTTCCATCCTTTGGCAAAATTAATTAACATGTTT - -GGAATTA	2096
QY	2251	GGATCCCTGAAGAAAACGACAGAAATAGTAACATGTTTCTGTATTTGTTTCTAGTCCTGG	2310
Db	2097	TGATTAACACCATTAATAAGCATGATGACGAATTAATTAATTCATGTAATTCGTATTTGGG	2156
QY	2311	AATTAATTTCTTTTATTAACATTTTTCCTCCAGGCGTCCACATTTGGCAACGCTGGGAGAT	2370
Db	2157	TGTTATTTGGCTTTGTGCACTTTATTCATGCAAGGATTAATTTAAGCGCAGACAGGGAAAT	2216
QY	2371	CCTCACTAAGGGGCTGCATATCATAGTGGTTTCCAGATCTCATGCTGACAGAGATGCAGCTG	2430
Db	2217	TTTAACGATGAGATTTAAGACACTTGGCCCTTCAAGCCCATGTTTATACAGATATATGCTGTG	2276
QY	2431	GTTTGATGACCTTAATAAACACCACTGGAGCATTGACAAACAGGCTTGGCAATATGCGGC	2490
Db	2277	GTTTGATGAAAAAGAAAAACAGCACAGGAGGCTTGACAACAATTAATGCCATAGATTAAC	2336
QY	2491	TCAAGTTAAAGGGGCTATAGGTTCCAGGCTTGCTGTCATTAACCAATATATGCAAAATCT	2550
Db	2337	ACAAATTCAGAGACACAGSTTCCAGGATTTGGCGCTTAACACAAAATGCAACTAACAT	2396
QY	2551	TGGGACAGGCATTAATTAATCCTTAATCTATGTTGGCAATTACACTTTACTCTTAAGC	2610
Db	2397	GGGACTTTCAGTTATATCAATTTCCCTTTATATATGATGGAGATGACATTCCTATTTCTGAG	2456
QY	2611	AATGTACCCATATGATGCAATACAGAGAGTGTGTAATGAATAATGTTGCTGGACAAC	2670
Db	2457	TATTTGCTCCAGTACTTGCCTGACAGGAATGATGAACCCGACGAATGACTGATTTTGC	2516
QY	2671	ACTGAAAGATTAAGAAAGAGCTAGAAAGAGACTGGAAGATTTGCTACAGAAAGCATCGAAAA	2730
Db	2517	CAACAAAGATTAACACAGAACTTAAGCATGTGGAAGAATAGCACTGAAGCTTTGGAGAA	2576
QY	2731	CTTCCGAAAGCTTGTCTTTTACATCGGAGCAGAAATTTGAATATACATGATGACAGAG	2790
Db	2577	TATACGTACTATAGTCTCTTAATAACAGGAAAAAACCTTCAGCAAAATGTATGAAGAT	2636
QY	2791	TTTCAAGTACCATACAGAAATCTTTAGAGAAAGCAACACTTTCGGGGGCTCATATTTTC	2850
Db	2637	GCCTGAGACTCAACACAGAAATACCTCGAAGAAAGACACAGATTAATGGAAGCTTTATGTC	2696
QY	2851	TATCACCCAGGCAATGATGATTTTTCCTATGCTGCGCTGTTTCCGTTTGGTGGTGCATCT	2910
Db	2697	ATTTCAGCATTGCTTTATATATATTTTGGCTATGACAGAGGGTTTCGATTTGGAGCCATTTT	2756
QY	2911	GGTGCAATGATGTTTATGAACTTTCAGGATGTCTTTTGGTATTTCCAGCTATATGCTTT	2970
Db	2757	AATTCAGCTGAGCAATGACACCCAGAGGGCATGTATATGTTTATACGCAATTCATATA	2816
QY	2971	TGCTGCCATGCAAGTGGGGCAGGTCAAGTCAATTTTGTCTCGACTATAGCCAAACCAAT	3030
Db	2817	TGGACCTATGCGCAATCGAAAAAAGCGTCTGTTTGGCTCGTGAATATTCCAAAGCCAAATC	2876
QY	3031	ATCAGACAGCCCACTCATATGATATCATATGAAAAAGCCCTCTGATATGACAGCTACAGCCC	3090
Db	2877	GGGGGCTGCGCATATGTTTGTCTTGTGGAAAAAGAAACCAATATATGACAGCGCGAGTCA	2936

GenCore version 5.1.3
Copyright (c) 1993 - 2002 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: December 9, 2002, 17:09:38 ; Search time 29467 Seconds

(without alignments)
2351.800 Million cell updates/sec

Title: US-09-672-725C-1

Perfect score: 4279
Sequence: 1 ggaagcagagatcgagatg.....caaaaaaaaaaaaaaa 4279

Scoring table: IDENTITY_NUC
Gapop 10.0 , Gapext 1.0

Searched: 16154066 seqs, 8097743376 residues

Total number of hits satisfying chosen parameters: 32308132

Minimum DB seq length: 0

Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database :

EST:*
1: em_estba:*
2: em_esthum:*
3: em_estin:*
4: em_estnu:*
5: em_estcov:*
6: em_estpl:*
7: em_estro:*
8: em_hlc:*
9: gb_est1:*
10: gb_est2:*
11: gb_hlc:*
12: gb_est3:*
13: gb_est4:*
14: gb_est5:*
15: em_estfun:*
16: em_estom:*
17: gb_gss:*
18: em_gss_hum:*
19: em_gss_inv:*
20: em_gss_pla:*
21: em_gss_vrt:*
22: em_gss_fun:*
23: em_gss_mam:*
24: em_gss_mus:*
25: em_gss_other:*
26: em_gss_pro:*
27: em_gss_tod:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	769.8	18.0	943	AL520322	AL520322 AL520322
2	738	17.2	929	14	BO882401
3	723.8	16.9	998	14	BM904842
4	651.4	15.2	939	14	BM904842
5	612.6	14.3	861	14	BO720763
6	576.8	13.5	981	12	BO717101
					BF969667
					BF969667 602272046

C	7	520.8	12.2	643	14	BO366012
	8	518.6	12.2	785	10	AV709991
	9	518.6	12.1	851	12	BF969062
	10	516.4	12.1	1019	12	BF969062
	11	510.6	11.9	894	12	BF969062
	12	509.6	11.9	803	12	BF969062
	13	479	11.2	587	10	BM053000
	14	479	11.2	605	10	BM053000
	15	475.2	11.1	795	12	BF969062
	16	467	10.9	563	13	BF969062
	17	467	10.9	563	13	BF969062
	18	467	10.9	563	13	BF969062
	19	463.8	10.8	563	13	BF969062
	20	449.4	10.5	544	12	BF969062
	21	448	10.5	760	9	AL520321
	22	438.4	10.2	726	12	BF969062
	23	436.4	10.2	546	13	BF969062
	24	417.4	9.8	662	13	BF969062
	25	408	9.5	759	10	BF969062
	26	407.4	9.5	517	13	BF969062
	27	386.2	9.0	598	9	AL807162
	28	386	9.0	1201	12	BF969062
	29	383.8	9.0	511	9	AA239727
	30	373.8	8.7	560	12	BF969062
	31	372	8.7	482	13	BF969062
	32	368.2	8.6	584	13	BF969062
	33	358	8.4	609	10	AM188349
	34	356	8.3	519	9	AA434959
	35	353.2	8.3	431	14	BO311985
	36	352.8	8.2	528	9	AA456377
	37	349.4	8.2	713	9	AT765301
	38	337.8	7.9	414	12	BF75493
	39	334.2	7.8	411	13	BM432014
	40	333.8	7.8	478	10	BE217416
	41	329	7.7	628	12	BM080311
	42	328.6	7.7	890	12	BM105127
	43	327.6	7.7	888	9	AV287606
	44	325.4	7.6	822	13	BI332761
	45	325	7.6	453	13	BM658768

ALIGNMENTS

RESULT 1
LOCUS AL520322 943 bp mRNA linear EST 13-FEB-2001
DEFINITION AL520322 LTI_NFL004_NBC2 Homo sapiens CDNA clone CS0DB006Y15 5
ACCESSION AL520322
VERSION AL520322.1 GI:12783815
KEYWORDS EST.
SOURCE human.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1 (bases 1 to 943)
AUTHORS Li, W.B., Gruber, C., Jesse, J., and Polayes, D.
TITLE Full-length CDNA libraries and normalization
JOURNAL Unpublished (2001)
COMMENT Contact: Genoscope
Genoscope - Centre National de Sequencage
BP 191 91006 EVRY cedex - France
Email: segre@genoscope.cns.fr, Web : www.genoscope.cns.fr.

FEATURES

source
1..943
/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone="CS0DB006Y15"
/clone_lib="LTI_NFL004_NBC2"
/sex="male"
/tissue_type="neuroblastoma cells"
/lab_host="DH10B"

/note="Organ: brain; Vector: PCMVSPORT 6; 1st strand cDNA was primed with a NotI-oligo(dT) primer. Five prime end enriched, double-stranded cDNA was digested with Not I and cloned into the Not I and Eco RV sites of the PCMVSPORT 6 vector. Library was normalized. Library was constructed by Life Technologies. Contact: Feng Liang Life Technologies, a division of Invitrogen 9800 Medical Center Drive, Rockville, Maryland 20850, USA Fax: (1) 301 610 8371 Email: fliang@life.com URL: http://fulllength.invitrogen.com"

BASE COUNT 253 a 225 c 253 g 211 t 1 others
ORIGIN

Query Match 18.0%; Score 769.8; DB 9; Length 943;
Best Local Similarity 89.8%; Pred. No. 4.1e-123;
Matches 847; Conservative 1; Mismatches 93; Indels 2; Gaps 2;

QY 2944 TCTTTGATCTCAGCTATTGTCTTGGTGGCCATGGCAGGGGAGTTCATT 3003
DB 1 TCTGTATATTTTTCAGCTGTGTCTTGTGTCATGGCCGGGCGAAGTCATT 60
QY 3004 TGGCCGAGATATGCCAAAGCCAAAGTATCAGACGCCAGTCATCATATTGAAA 3063
DB 61 TGTCTCTGACTATGCCAAAGCCAAATATCAGACGCCAGTCATCATATTGAAA 120
QY 3064 AAGCCCTCTGATGACAGCTACAGCCCTCAGGGCCCTCAAGCCAAATAGCTGGAAGAAA 3123
DB 121 AAGCCCTTGTATGACAGCTACAGCCAGGAGGCTTAATGCCGAACACTTGAAGAAA 180
QY 3124 TGTGCAATTAATGAGTCTGTTTCAACTATCCACTGACAGACATATCCCGTCTCCA 3183
DB 181 TGTCACTATTTGATGATGATGATATCCACCCAGCCGACATCCAGTCTCTCCA 240
QY 3184 GGGGCTGAGCCTCGAGCTGGAAGAGGGCCAGAGCTGGCCCTGATGATGAGTGGCTG 3243
DB 241 GGGACTGAGCCTGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 300
QY 3244 TGGGAAGAGCAGTGTGTTCACTCTAGAGCGCTCTATGACCCCTGGTGTGTTCACT 3303
DB 301 TGGGAAGAGCAGTGTGTTCACTCTAGAGCGCTCTATGACCCCTGGTGTGTTCACT 360
QY 3304 GCTAATTTGATGCAAGAGATTAAGCACTGATGTCAGTGGCTCGAGACACCTGGG 3363
DB 361 GCTGCTTGTATGCAAGAGATTAAGCACTGATGTCAGTGGCTCGAGACACCTGGG 420
QY 3364 CATGCTGTCTCAGAGAGCCATCTCTTGTACTGACGATTTGCCGAGAACTTGCTATGG 3423
DB 421 CATGCTGTCTCAGAGAGCCATCTCTTGTACTGACGATTTGCCGAGAACTTGCTATGG 480
QY 3424 AGACAAACAGCCGGGTGTTCACTAGAGAGATTAATGAGGAGCAGCAGGAGGCAACAT 3483
DB 481 AGACAAACAGCCGGGTGTTCACTAGAGAGATTAATGAGGAGCAGCAGGAGGCAACAT 540
QY 3484 ACACCACTTATGATGAGACACTCCCTGAGAAATATCAACACAGATGAGAGCAAGAGAC 3543
DB 541 ACATGCTTATGATGAGACACTCCCTGAGAAATATCAACAGATGAGAGCAAGAGAC 600
QY 3544 CCAGCTCTGTGTGGCCAGAGAAAGAGGCAATGCAATGCTGCTCTTGTAGACAGCC 3603
DB 601 TCAGCTCTGTGTGGCCAGAGAAAGAGGCAATGCAATGCTGCTCTTGTAGACAGCC 660
QY 3604 TCATATTTTGTGTTGGATGAGACTACAGCTCTGATGATGAGAAAGTGAAGAGTTGT 3663
DB 661 TCATATTTTGTGTTGGATGAGAGCAAGCTAGCTCTGATGATGAGAAAGTGAAGAGTTGT 720
QY 3664 CCAGAGAGCCCTGAGCAAAAGCCAGAGAGCCGACCTGATGATGATGATGATGATGAT 3723
DB 721 CCAGAGAGCCCTGAGCAAAAGCCAGAGAGCCGACCTGATGATGATGATGATGATGAT 780
QY 3724 GTCCACCATTCAGAGATGAGATTAATAGTG-GTGTTCAGATGAGCAAGTCAAGAGC 3782
DB 781 GTCCACCATTCAGAGATGAGATTAATAGTG-TGTGTGTGTGTGTGTGTGTGTGTGTGT 840

QY 3783 ATGGCACATCATCAGAGCTGCTGGCCAGAAAGCATCTATTTTCCATGTCAGTGTCC 3842
DB 841 ATGGCACATCATCAGAGCTGCTGGCCAGAAAGCATCTATTTTCAATGTCAGTGTCC 900
QY 3843 AGGCTGAGCAAAAGCCGTGTGATGATGATGATGATGATGATGATGATGATGATGAT 3885
DB 901 AGGCTGAGCAAAAGCCG-AGTGAATCTGATGATGATGATGATGATGATGATGATGAT 942

RESULT 2

BQ882401 929 bp mRNA linear EST 16-AUG-2002
LOCUS AGENCOURT_8627902 NIH_MGC_43 Homo sapiens cDNA IMAGE:6291782
DEFINITION 5' mRNA sequence.
ACCESSION BQ882401
VERSION BQ882401.1 GI:22274409
KEYWORDS EST.
SOURCE human.
ORGANISM Homo sapiens
Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
1 (bases 1 to 929)
NIH-MGC http://mgi.nci.nih.gov/
National Institutes of Health, Mammalian Gene Collection (MGC)
Unpublished (1999)
Contact: Robert Strausberg, Ph.D.
Email: cgabs@emill.nih.gov
Tissue Procurement: ATCC
CDNA Library Preparation: Rubin Laboratory
CDNA Library Arrayed by: The I.M.A.G.E. Consortium (LNL)
DNA Sequencing by: Agencourt Bioscience Corporation
Clone distribution: MGC clone distribution information can be
found through the I.M.A.G.E. Consortium/LNL at:
http://image.lnl.gov
Plate: LNCM2492 row: C column: 15
High quality sequence stop: 677.
Location/Qualifiers

FEATURES

1..929
/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone="IMAGE:6291782"
/clone_lib="NIH_MGC_43"
/tissue_type="normal pigmented retinal epithelium"
/lab_host="DH10B (phage-resistant) site:1; xhoi: site:2;
/note="Organ: eye; Vector: pORF7; Site:1: xhoi: site:2;
EcoRI; CDNA made by oligo-dT priming. Directionally
cloned into EcoRI/XhoI sites using the following 5'
adaptor: GGCACGAC(G). Library constructed by Ling Hong
in the Laboratory of Gerald M. Rubin (University of
California, Berkeley) using ZAP-CDNA synthesis kit
(Stratagene) and Superscript II RT (Life Technologies).
Note: this is a NIH_MGC Library. 1 others

BASE COUNT 281 a 176 c 237 g 234 t 1 others
ORIGIN

Query Match 17.2%; Score 738; DB 14; Length 929;
Best Local Similarity 90.0%; Pred. No. 1.2e-117;
Matches 836; Conservative 0; Mismatches 86; Indels 7; Gaps 4;

QY 758 GCTATGAAAAGCTGGAGAGTAGTGAAGAGCTTAGCAATCAACT-GTAT 816
DB 1 GCGTATGAAAAGCTGGAGAGTAGTGAAGAGCTTAGCAATCAACTGCTGTAT 60
QY 817 TGCCTTGGAGAGCAAAAGAAAGACTTGAAGAGTACAAATAATTTGAGAGAGCTAA 876
DB 61 TGCATTTGGAGAGCAAAAGAAAGACTTGAAGAGTACAAATAATTTGAGAGAGCTAA 120
QY 877 AGGATTTGGAGAGAAAGAGTATCAGAGCCCAATTTCTATTTGTCGCTTTCTTAT 936
DB 121 AAGAAATTTGGAGAGAAAGAGTATTTACAGCCCAATTTCTATTTGTCGCTTTCTTAT 180
QY 937 GATCATGATATATGCTGCTGCTTCTGATATGAGACTCTTGCTCTCCAGTGA 996

Db 779 GAGTGTATTAACCTATCCACCCGACGACATCCAGTCTTACGGAGCTAGCCCTG 838
Oy 3197 GAGGTGAAGAAGGCGCAGACGCTGCCCTGT AGTAGACGTGGTGTGGGAAGACAC 3255
Db 839 GAGGTGAAGAAGGCGACACCTGTCTGTGGGCGACAGTGGCTGGGAAGACAC 898
Oy 3256 AGT-TGTTCACTCTCTTATGAGGCTTATGACCCCTGGCTGGTGCATGCTAATTGAT 3314
Db 899 AGTGGTCCACTCTCTGAGGCGGTCTACACCCCTTGGCAAGAAATGCTCTTGATG 958
Oy 3315 GCAAGAGATAAAGC 3329
Db 959 GCAAAAGATAAAGC 973

RESULT 4
B0720763 939 bp mRNA linear EST 16-JUL-2002
LOCUS B0720763
DEFINITION AGENCOURT_8233170 Lupski_dorsal_root_ganglion Homo sapiens cDNA
ACCESSION B0720763
VERSION B0720763.1 GI:21859660
KEYWORDS EST.
SOURCE human.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominiidae; Homo.
REFERENCE 1 (bases 1 to 939)
AUTHORS NIH-MGC http://mgc.nci.nih.gov/
TITLE National Institutes of Health, Mammalian Gene Collection (MGC)
JOURNAL Unpublished (1999)
COMMENT Contact: Robert Strausberg, Ph.D.
Email: cgabps-remail.nih.gov
Tissue Procurement: Dr. James R. Lupski
cDNA Library Preparation: Life Technologies, Inc.
DNA Library Arrayed by: The I.M.A.G.E. Consortium (LNLN)
CDNA Sequencing by: Agencourt Bioscience Corporation
Clone distribution: MGC clone distribution information can be
found through the I.M.A.G.E. Consortium/LNLN at:
http://image.llnl.gov
plate: LLM13573 row: 0 column: 22
High quality sequence stop: 646.

FEATURES
source
location/Qualifiers
1. 939
/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone_image="6184557"
/clone_lib="Lupski_dorsal_root_ganglion"
/sex="male"
/tissue_type="dorsal root ganglia"
/dev_stage="adult, 36 yr"
/lab_host="DH10B"
/note="Vector: PCMV-SPORT6 (Life Technologies); Site_1:
NotI; Site_2: SalI; cDNA made by oligo-dT priming.
Directionally cloned using the following adaptors:
5'-TCGACCCAGGCGTCG-3' and
5'-GACTAGTCTTAGATTCGAGCGGCGCCCT(15)-3'. Size selected >
1 kb for average insert length 1.7 kb. This is a primary
library, non-amplified. Library constructed by Life
Technologies and donated by J. Lupski, M.D./Ph.D. (Baylor
College of Medicine) and is available through Life
Technologies."

BASE COUNT 272 a 202 c 221 g 244 t
ORIGIN
Query Match 15.2%; Score 651.4; DB 14; Length 939;
Best Local Similarity 91.3%; Pred. No. 1.e-102;
Matches 713; Conservative 0; Mismatches 66; Indels 2; Gaps 2;

Oy 2335 GCAAAAGCTGGAGATCTCTCACTAAGGCGCTGATCATGTTTTCAGATCATGCTGA 2414
Db 1 GCAAAAGCTGGAGATCTCTCACTAAGGCGCTGATCATGTTTTCAGATCATGCTCA 60

Oy 2415 GACAGATGTCAGCTGTTTGATGACCTTAAAAACACACACTGAGCATGACACAGGC 2474
Db 61 GACAGATGTCAGCTGTTTGATGACCTTAAAAACACACACTGAGCATGACATACAGGC 120
Oy 2475 TTGCAATGATGCGGCTCAAGTTAAAGGGCTATAGGTTCCAGGCTTGTCTGATTAACC 2534
Db 121 TCGCCATGATGCTGCTCAAGTTAAAGGGCTATAGGTTCCAGGCTTGTCTGATTAACC 180
Oy 2535 AGAATATAGCAAACTTGGGACAGGCACTTATATCCCTTAATCTATGTTGGCAATTAA 2594
Db 181 AGAATATAGCAAACTTGGGACAGGCACTTATATCCCTTAATCTATGTTGGCAATTAA 240
Oy 2595 CACTTTTACCTCTTGAATTTTACCTATCCATTCATAGCAGAGATGTTGAATGAAA 2654
Db 241 CACTTTTACCTCTTGAATTTTACCTATCCATTCATAGCAGAGATGTTGAATGAAA 300
Oy 2655 TGTGTCGACACAGCACTGTAAGATAGAAAGCTAGAAAGAGCTGGGAATGCTA 2714
Db 301 TGTGTCGACACAGCACTGTAAGATAGAAAGCTAGAAAGAGCTGGGAATGCTA 360
Oy 2715 CAGAAAGCATGGAACCTCGGAACCTGTTCTTGTGACTCGGAGCAGAGTTGAAAT 2774
Db 361 CAGAAAGCATGGAACCTCGGAACCTGTTCTTGTGACTCGGAGCAGAGTTGAAAT 420
Oy 2775 ACATGTATGACACAGAGTTTGCATACCATACAGAAACTTTGAGAAAGCACATCT 2834
Db 421 ACATGTATGACACAGAGTTTGCATACCATACAGAAACTTTGAGAAAGCACATCT 480
Oy 2835 TCGGGTCTCATTTTCTATACCCGCAAGATGATTTTCTATGCTGCTGTTCC 2894
Db 481 TTGGAATTCATTTTCTTCCACCGCAAGATGATTTTCTATGCTGCTGTTCC 540
Oy 2895 GGTGTCCTACTTGTGGTGGCAATGATCATGAACTTTCAGAGTTCTTTGGTAT 2954
Db 541 GGTGTCCTACTTGTGGTGGCAATGATCATGAACTTTCAGAGTTCTTTGGTAT 600
Oy 2955 TCTCAGCTATTTGTTGTTGTCATGCGAGAGTGGGCTAGTTCCTCTGACT 3014
Db 601 TTTCACTGTTGCTTGTGTCATGCGAGCGTGGGCAAGTCAATTTCTCTGACT 660
Oy 3015 ATG-CCAAAGCCAAATGATCAGAGCCGACGATCATGATGATGAAAAGCCCTG 3073
Db 661 ATG-CCAAAGCCAAATGATCAGAGCCGACGATCATGATGATGAAAAGCCCTG 720
Oy 3074 ATTGACAGCTACAGC-CCCTACGCGCTCAAGCCAAATACGTTGGAAGAATGACAT 3132
Db 721 ATTGACAGCTACAGCAGCAGGAGCCCTAATGCCGAACATTTGGAAGAATGACAT 780
Oy 3133 T 3133
Db 781 T 781

RESULT 5
B0717101 861 bp mRNA linear EST 16-JUL-2002
LOCUS B0717101
DEFINITION AGENCOURT_8101135 Lupski_sympathetic_trunk Homo sapiens cDNA clone
IMAGE:6189786 5', mRNA sequence.
B0717101
ACCESSION B0717101 GI:2185998
VERSION B0717101.1
KEYWORDS EST.
SOURCE human.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominiidae; Homo.
REFERENCE 1 (bases 1 to 861)
AUTHORS NIH-MGC http://mgc.nci.nih.gov/
TITLE National Institutes of Health, Mammalian Gene Collection (MGC)
JOURNAL Unpublished (1999)
COMMENT Contact: Robert Strausberg, Ph.D.
Email: cgabps-remail.nih.gov
Tissue Procurement: Dr. James R. Lupski

cDNA Library Preparation: Life Technologies, Inc.
cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LINTL)
DNA Sequencing by: Agencourt Bioscience Corporation
Clone distribution: MGC clone distribution information can be
found through the I.M.A.G.E. Consortium/LINTL at:
<http://image.llnl.gov>
plate: L1M13587 row: 1 column: 19

High quality sequence stop: 603.
Location/Qualifiers

FEATURES

source

```
1. 861
/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone="IMAGE:6189786"
/clone_lib="Lupski_symphatic_trunk"
/sex="male"
/tissue_type="sympathetic trunk"
/dev_stage="adult, 16 yr"
/lab_host="DH10B"
/Note="Vector: PCMV-SPORT6 (Life Technologies); Site_1:
NotI; Site_2: SalI; cDNA made by oligo-dT priming:
Directionally cloned using the following adaptors:
5'-TCGACCCACGCGCCG-3' and
5'-GACTAGTCTAGATCGCAGGCGGCCCT(15)-3'. Size selected >
1 kb for average insert length 1.9 kb. This is a primary
library, non-amplified. Library constructed by Life
Technologies and donated by J. Lupski, M.D./Ph.D. (Baylor
College of Medicine); available through Life
Technologies."
```

BASE COUNT

236 a 191 c 216 g 218 t

Query Match 14.3%; Score 612.6; DB 14; Length 861;
Best Local Similarity 88.3%; Pred. No. 5.4e-96;
Matches 688; Conservative 0; Mismatches 89; Indels 2; Gaps 2;

```
QY 2586 GGCATTAACTTTTACTCTTGGCAATTTGACCATTCATTCGCAATAGCAGGAGTTGTG 2645
DB 1 GGCACCTAACACTGTTACTCTTAGCAATTTGACCATTCATTCGCAATAGCAGGAGTTGTG 60
QY 2646 AATGTAATGTTGTCGTGACACACACACACACACACACACACACACACACACACACAC 2705
DB 61 AATGTAATGTTGTCGTGACACACACACACACACACACACACACACACACACACACAC 120
QY 2706 AGATTGCTACAGAGCCATCGAAACTTCGCACTGTTGTTCTTTGACCTCGGACACGA 2765
DB 121 AGATGCTACTGAGCAATAGAAAACCTCCGAACCGTTGTTCTTTGACCTCGGACACGA 180
QY 2766 AGTTGGAATACATGATGACACAGAGTTTGCAGTACCATACAGAACTCTTTGAGGAAG 2825
DB 181 AGTTGGAATACATGATGACACAGAGTTTGCAGTACCATACAGAACTCTTTGAGGAAG 240
QY 2826 CACACATCTCGGGGCTCTCATTTTCTATCCAGGCAATGATGATTTTTCCTATGCTG 2885
DB 241 CACACATCTCGGGGCTCTCATTTTCTATCCAGGCAATGATGATTTTTCCTATGCTG 300
QY 2886 GCTGTTCCGTTTGGTGGCTACTTGTGTGCAATAGTTTCATGAACTTCAGSAGTTTC 2945
DB 301 GATGTTCCGTTTGGTGGCTACTTGTGTGCAATAGTTTCATGAACTTCAGSAGTTTC 360
QY 2946 TTTTGGTATCTCAGCATTTGCTTGGTGGCAATGCGAGTGGGCAAGTCACTTCAATTG 3005
DB 361 TGTAGATTTTTCAGCTGTTGTTGTTGGTGGCAATGCGAGTGGGCAAGTCACTTCAATTG 420
QY 3006 CTCTGACTATGCGCAAGCAAGTATGACAGCCAGCTCATGATGATGATGAAAAA 3065
DB 421 CTCTGACTATGCGCAAGCAAGTATGACAGCCAGCTCATGATGATGATGAAAAA 480
QY 3066 GCCCTGTGATGAGCTACAGCTCAGAGGCTTAAGCCAAATAGCTTTGAGGAAGATG 3125
DB 481 CCCCTTGTATGACAGCTACAGAGGCTTAATGCCAATGATTTGAGGAAGATG 540
QY 3126 TGCATTTT AATGAGGTGTTTCAATCATCTGACACAGCAATCCCGTCTCCAG 3184
DB 1 TGCATTTT AATGAGGTGTTTCAATCATCTGACACAGCAATCCCGTCTCCAG 3184
```

```
DB 541 TCACATTTGGTGAAGTTGATTCACATATCCACCCGACCGACATPCCAGTGTTCAG 600
QY 3185 GGGCTGACCTTCGAGGTGAGAGAGGCGCCAGACGCTGCGCTCTAGTAGCAGTGTCTGT 3244
DB 601 GGACTGAGCCCTGGAAGTGAAGAGGGGACAGACCTGCTGTGTGGGACAGAGTGTCTGT 660
QY 3245 GGAAGAGCAGATGTTGCTGCTCTGAGAGGCTCTCTATGACCCCTTGGCTGTTCAGTG 3304
DB 661 GGAAGAGCAGATGTTGCTGCTCTGAGAGGCTCTCTATGACCCCTTGGCAGGAAAGTG 720
QY 3305 CTAATTTATGCGCAAGAGATTAAGACACCTGAATGTCCAGTGTCCGACACACCTGGG 3363
DB 721 CTGCTGATGCGCAA AATTAAGCGACGTGAAGTTCAATTGGCTCGAACCCTCTGG 778
```

RESULT 6
BF969667 981 bp mRNA linear EST 22-JAN-2001
LOCUS 602272046F1 NIH_MGC_84 Homo sapiens cDNA clone IMAGE:4360090 5',
DEFINITION mRNA sequence.
ACCESSION BF969667
VERSION BF969667.1 GI:12336882
KEYWORDS EST.
SOURCE human.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
REFERENCE NIH-MGC <http://mgc.nci.nih.gov/>.
Tissue Procurement: ATCC
Contact: Robert Strausberg, Ph.D.
Email: cgabbs-remail.nih.gov
Unpublished (1999)
National Institutes of Health, Mammalian Gene Collection (MGC)

COMMENT
cDNA Library Preparation: Life Technologies, Inc.
cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LINTL)
DNA Sequencing by: Incyte Genomics, Inc.
Clone distribution: MGC clone distribution information can be
found through the I.M.A.G.E. Consortium/LINTL at:
<http://image.llnl.gov>
plate: L1M10000 row: 1 column: 11
High quality sequence stop: 670.
Location/Qualifiers

FEATURES

source

```
1. 981
/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone="IMAGE:4360090"
/clone_lib="NIH_MGC_84"
/tissue_type="adrenal cortex carcinoma, cell line"
/lab_host="DH10B (phage-resistant)"
/Note="Organ: adrenal gland; Vector: PCMV-SPORT6; Site_1:
NotI; Site_2: SalI; Cloned unidirectionally; oligo-dT
primed. Average insert size 1.229 kb. Library enriched for
full-length clones and constructed by Life Technologies.
Note: this is a NIH-MGC Library."
```

BASE COUNT 267 a 194 c 250 g 270 t

Query Match 13.5%; Score 576.8; DB 12; Length 981;
Best Local Similarity 92.1%; Pred. No. 7.6e-90;
Matches 630; Conservative 0; Mismatches 52; Indels 2; Gaps 2;

```
QY 2393 ATGCTTTTTCAGATTCATGCTGAGACAGATGTCACCTGTTGATGACCTTAAGAACAC 2452
DB 1 ATGCTTTTTCAGATTCATGCTGAGACAGATGTCACCTGTTGATGACCTTAAGAACAC 60
QY 2453 ACTGAGAGCAATTACCAACAGGCTTCCCAATGATGCGGCTCAAGTTTAAAGGGCTATAGT 2512
DB 61 ACTGAGAGCAATTACCAACAGGCTTCCCAATGATGCGGCTCAAGTTTAAAGGGCTATAGT 120
QY 2513 TCCAGGCTTGTGCTATTAACCAAGATATAGCAATCTTGGACAGCATTTATATTC 2572
DB 121 TCCAGGCTTGTGCTATTAACCAAGATATAGCAATCTTGGACAGCATTTATATATTC 180
```


CACTGAAGATTAAGAAAGAGCTGAGAAGGAGCTGGGAAGATTGCCCTACTGACCGCCGCGC... 2670
.....CCCCCCCCCCGCGTTCGTTTGCTGTGGACCAAG 180


```

|||||
Db 181 CACGAAAGTAAGAAAGACTGAAAGGTCTGGGAAAGATCGCTACTGTAAGCAATAGAAA 240
Oy 2730 ACTCCGAACTGTTGTTCTTTTCTGAGGAGAGAAAGTTTGAAATCATGTATGCACAGA 2789
Db 241 ACTCCGAACTGTTGTTCTTTTCTGAGGAGAGAAAGTTTGAAATCATGTATGCACAGA 300
Oy 2790 GTTTCGAATGACATACAGAAAGCTTTGAGGAAAGACATCTCTGGGGTCTCATTTT 2849
Db 301 GTTTCGAATGACATACAGAAAGCTTTGAGGAAAGACATCTCTGGGGTCTCATTTT 360
Oy 2850 CTATCACCAGGCAATGATGATTTTCTATGCTGGCTGTTCCGGTTGGTCTTACT 2909
Db 361 CTTTACCAGGCAATGATGATTTTCTATGCTGGCTGTTCCGGTTGGTCTTACT 420
Oy 2910 TGTGGCAATGATGATGATTTTCTATGCTGGCTGTTCCGGTTGGTCTTACT 2969
Db 421 TGTGGCAATGATGATGATTTTCTATGCTGGCTGTTCCGGTTGGTCTTACT 480
Oy 2970 TGTGGCAATGATGATGATTTTCTATGCTGGCTGTTCCGGTTGGTCTTACT 3028
Db 481 TGTGGCAATGATGATGATTTTCTATGCTGGCTGTTCCGGTTGGTCTTACT 540
Oy 3029 GTATCAGACGCCCATCATCATGATCATTTGAAAAAGCCCTCGATTGACAGTACAGC 3088
Db 541 ATATCAGACGCCCATCATCATGATCATTTGAAAAAGCCCTCGATTGACAGTACAGC 599
Oy 3089 CCTCAGCGCTCAGAGCC---AAATACGTTGAGAGAAATGACATTTAATGAGTCTG 3145
Db 600 ACGAAGGCTTATTTGCGGGAACATTTGGAGGAAATGACATTTGAGTCTG 659
Oy 3146 TTCACTATCCACATC-----GACCAACATCCCGCTGCTCAGGAGGCTGAGCTCGA 3198
Db 660 GTTTTCAATTTATTCACACCGTAGCGAGTCCAGTCCAGTCTCAGGAGTGAAGCTGG 719
Oy 3199 GGTGAAGAGGCGCAGAGCGTGGCCCTG 3227
Db 720 GTGAAGAAAGGCCAAAGCGGCTCTCG 748

RESULT 10
Bg248052 1019 bp mRNA linear EST 13-FEB-2001
LOCUS 60235987F1 NCL_CGAP_Mam1 Mus musculus cDNA clone IMAGE:4488404 5',
DEFINITION mRNA sequence.
ACCESSION Bg248052
VERSION Bg248052.1 GI:12757867
KEYWORDS house mouse.
SOURCE Mus musculus
ORGANISM Mus musculus; Chordata; Craniata; Vertebrata; Euteleostomi;
Eukaryota; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
REFERENCE NIH-MGC http://mgi.nci.nih.gov/
AUTHORS National Institutes of Health, Mammalian Gene Collection (MGC)
TITLE Unpublished (1999)
JOURNAL Contact: Robert Strausberg, Ph.D.
COMMENT Email: egabbs-remail.nih.gov
Tissue Procurement: Gilbert Smith, Ph.D.
cDNA Library Preparation: Life Technologies, Inc.
CDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
DNA Sequencing by: Incyte Genomics, Inc.
Clone distribution: MGC clone distribution information can be
found through the I.M.A.G.E. Consortium/LLNL at:
http://image.llnl.gov
Plate: LLAM10334 row: n column: 21
High quality sequence stop: 650.
Location/Qualifiers
1..1019
/organism="Mus musculus"
/strain="FVB/N"
/db_xref="taxon:10090"
/clone="IMAGE:4488404"

```

```

BASE COUNT 276 a 262 c 283 g 196 t 2 others
ORIGIN
Query Match 12.1%; Score 516.4; DB 12: Length 1019;
Best Local Similarity 83.5%; Pred. No. 2e-79;
Matches 611; Conservative 0; Mismatches 116; Indels 5; Gaps 2;
/clone.lib="NCL_CGAP_Mam1"
/tissue_type="tumor, biopsy sample"
/dev_stage="10 months, virgin"
/lab_host="DH10B"
/notes="Organ: mammary; Vector: pCMV-Sport6; Site:1: SalI;
Site:2: NotI; Cloned unidirectionally. Primer: Oligo dt.
Library constructed by Life Technologies. Investigator
providing samples: Gilbert Smith, NIH"

Oy 3113 TTGGAAGAAATGTGACATTTAATGAGGCTGCTCAACTATCCACTGACAGACATC 3172
Db 1 TTGGAAGAAATGTGACATTTAATGAGGCTGCTCAACTATCCACTGACAGACATC 60
Oy 3173 CCCGCTCCAGAGGCTGAGCCTGAGGTGAAGAGGCGCCAGACCTGCGCTCTAGGT 3232
Db 61 CCAGTCTTCAGAGGCTGAGCCTGAGGTGAAGAGGCGCCAGACGCTGCGCTCTAG 120
Oy 3233 AGCAGTCTGCTGGAGAGACAGAGTGTTCAGCTCTCTAGAGCCTCTATGACCCCTTG 3292
Db 121 AGCAGTCTGCTGGAGAGACAGAGTGTTCAGCTCTCTAGAGCCTCTATGACCC 180
Oy 3293 GCTGCTCAGTCTATTTGATGAGCAAGAGATTAAGACCTGAATGCTCAGTGGCTCGA 3352
Db 181 GCTGCTCAGTCTCTATGATGAGCAAGAGATTAAGACCTGAATGCTCAGTGGCTCGA 240
Oy 3353 GCACACTGGGCTGCTGCTCAGAGGCCCATCTGTTTGGCTGCTGACATTCGAGAGAC 3412
Db 241 GCACACTGGGCTGCTGCTCAGAGGCCCATCTGTTTGGCTGCTGACATTCGAGAGAC 300
Oy 3413 ATTGCTATGAGAGACACAGCGGCTGCTATCAGATGAAGATTAATGACAGCCAG 3472
Db 301 ATTGCTATGAGAGACACAGCGGCTGCTGCTATGAGAGATTAATGAGAGCCAG 360
Oy 3473 GAGGCCAATACACCACTTCATGAGACACTCCCTGAGAAATACACCAAGAGTGA 3532
Db 361 GAGGCCAATACACCACTTCATGAGACACTCCCTGAGAAATACACCAAGAGTGA 420
Oy 3533 GACAAAGAACCCAGCTCTGCTGAGCAAGAACAGCCATGCACTGCTGCTCT 3592
Db 421 GACAAAGAACCCAGCTCTGCTGAGCAAGAACAGCCATGCACTGCTGCTCT 480
Oy 3593 GTTAGACAGCCTCATATTTTGTGTTGATGAGTACATCAGCTCTGATACAGAAAGT 3652
Db 481 GTTAGACAGCCTCATATTTTGTGTTGATGAGTACATCAGCTCTGATACAGAAAGT 540
Oy 3653 GAAAAGTTGCTCCAAAGACCCCTGAGCAAGAGGCGGCACTGCTGATGATC 3712
Db 541 GAAAAGTTGCTCCAAAGAGGCTGAGCAAGAGGCGGCACTGCTGATGATC 600
Oy 3713 GCCACCGCTGCTCCACATCAGAAATGCAATTAATAG--TGTGTTTCAGATGCA 3770
Db 601 GCTACCGCTGCTCCACATCAGAAATGCAATTAATAG--TGTGTTTCAGATGCA 660
Oy 3771 AAGTCAAGAGCATGACACATCAACAGC--TGTGCGCCAGAAAGCATCTATTTT 3827
Db 661 AAGTCAAGAGCATGACACATCAACAGC--TGTGCGCCAGAAAGCATCTATTTT 720
Oy 3828 CCATGCTCAGTG 3839
Db 721 CCAGGTGCTG 732

RESULT 11
Bg584668 894 bp mRNA linear EST 12-DEC-2000
LOCUS 602098406F1 NCL_CGAP_Co24 Mus musculus cDNA clone IMAGE:4218385 5',
DEFINITION mRNA sequence.

```

ACCESSION BF584668
 VERSION BF584668.1 GI:11658386
 KEYWORDS EST.
 SOURCE house mouse.
 ORGANISM Mus musculus
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
 1 (bases 1 to 894)
 NIH-MGC <http://mgs.nci.nih.gov/>
 National Institutes of Health, Mammalian Gene Collection (MGC)
 Unpublished (1999)
 Contact: Robert Strausberg, Ph.D.
 Email: cgabbs-remail.nih.gov
 Tissue Procurement: Jeffrey E. Green, M.D.
 cDNA Library Preparation: Life Technologies, Inc.
 DNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
 DNA Sequencing by: Incyte Genomics, Inc.
 Clone distribution: MGC clone distribution information can be
 found through the I.M.A.G.E. Consortium/LLNL at:
<http://image.llnl.gov>
 Plate: LHAM9798 row: 1 column: 02
 High quality sequence stop: 651.
 Location/Qualifiers
 1..894

FEATURES

source

/organism="Mus musculus"
 /strain="FVB/N"
 /db_xref="taxon:10090"
 /clone="IMAGE:4218385"
 /clone_lib="NCI CGAP Co24"
 /lab_host="DH10B (TI phage-resistant)"
 /note="Organ: colon; Vector: pCMV-Sport6; Site_1: NotI;
 Site_2: SalI; Cloned unidirectionally. Primer: Oligo dT.
 Average insert size 1.6 kb. Constructed by Life
 Technologies. Note: this is a NCI CGAP library."
 BASE COUNT 263 a 203 c 242 g 185 t 1 others

Query Match 11.9%; Score 510.6; DB 12; Length 894;
 Best Local Similarity 83.0%; Pred. No. 2.1e-78;
 Matches 629; Conservative 0; Mismatches 125; Indels 4; Gaps 4;

QY 1018 CACTCTCTCTTTCTGATTAATTTGGGGCTTTAGTATTGACAGAGCCCAAGCAT 1077
 DB 1 CACTCTCTCTTTCCGGTTAA-TGGAGCATTCAGTGTGGACAGGCATCTCAATAT 59
 QY 1078 TGAACATTTGCAACGAGAGGAGGAGCAGCTTATGAATCTTCAAGTAAATGACAATTA 1137
 DB 60 TGAACCTTCCCAATGACAGAGGAGCAGCTTATGAAGTCTTCAAAATATTAATTA 119
 QY 1138 ACCAAGCTTACAGCTATTGAGAGTGGACATTAACAGATTAATTAAGGAATTT 1197
 DB 120 GCCCAGTATACAGCTTCTCAAGAGTGGGCAAAACGACACATACAGGAATCT 179
 QY 1198 GGAATTCAAAATGTACCTTACCTTTCGAAAAGAGTTAGATCTTAAAGG 1257
 DB 180 GGAATTTAAGATTAATCTCACTTCACTTACCATCTCGAAAAAGATTGACATCTTGAAGG 239
 QY 1258 TCTCAACCTGAAGGTTGAGAGTGGGAGACAGAGGCGCTGGTGGAAAGTGGCGG 1317
 DB 240 CCTCAATCTGAAGGTGAAGAGCGGACAGAGGCGCTGGTGGCAAGGCGTGG 299
 QY 1318 GAAGAGACACACCGTGCAGTGTGAGAGGCTTATGAAGCCACAGATGAGCATGCTCG 1377
 DB 300 AAAAGACACAACTGTCCAGCTGATGCAAGGCTTACAGCCCTTATGAGCATGCTAG 359
 QY 1378 TATTTGATGACAGACATTAGGACCTAATTAAGGATCTTCGGAAATTTACTGTTG 1437
 DB 360 TATTCAGCGACAGACATTAGAACCATCAATGTGAGTATCTGAGGAGATCATTTGGTGT 419
 QY 1438 GGTGATTCAGAGCCTGTGTGTTGCCACACAGATAGCTGAAACATTTGCTTGGCG 1497
 DB 420 GGTGATTCAGAGCCTGTGTGTTGCCACACAGATAGCTGAAACATTTGCTTGGCG 479

QY 1498 GAAATGTCAACATGATGATTTGAGAAAGCTGTTAAGAAACCAATGCTATGATTT 1557
 DB 480 AGAATATCTCCATGATGATTTGAGAAAGCTGTTAAGAAACCAATGCTATGATTT 539
 QY 1558 TATCATGAATACCTATTAATTTGACACTGTGTTGAGAGAGAGGGGCCACGTGAG 1617
 DB 540 CATCATGAATGCCCCACCAATTTGACACTGTGTTGAGAGAGAGGGGCCACGTGAG 599
 QY 1618 TGTGTGACAGAAAGAAATGCGCATTTGCTCGGGCCCTGTTGCGACCCCAAGATCT 1677
 DB 600 TGGGGACAGAAACAGAAATCG-CATTTGCCCGGGCCCTGTTGCGAATCCCAAG-TGCC 657
 QY 1678 TCTGCTGATGAGGCAACGCTCAGCTGTGACACTGGAAGGAAGCAGTGTCTAGTGGC 1737
 DB 658 TTTGCTGAGAGAGCCACCTCAGC-CTGATATACAGAAAGTGAAGCTGGGGTTCAGGCCG 716
 QY 1738 CTTGATTAAGCCCAAGAAAGCCGCGACTACCATTTGGA 1775
 DB 717 ACCTAGTATAGGTATCAAAAGGGCGAGCACCACATTTGGA 754

RESULT 12

LOCUS

BF584668.1 NIH-MGC_94 Mus musculus cDNA clone IMAGE:451157 5',
 mRNA sequence.
 BG298756

ACCESSION BG298756
 VERSION BG298756.1 GI:13063728
 KEYWORDS EST.
 SOURCE house mouse.
 ORGANISM Mus musculus

REFERENCE Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
 1 (bases 1 to 803)
 NIH-MGC <http://mgs.nci.nih.gov/>
 National Institutes of Health, Mammalian Gene Collection (MGC)
 Unpublished (1999)
 Contact: Robert Strausberg, Ph.D.
 Email: cgabbs-remail.nih.gov
 Tissue Procurement: The Cepko Laboratory
 cDNA Library Preparation: Life Technologies, Inc.
 DNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
 DNA Sequencing by: Incyte Genomics, Inc.
 Clone distribution: MGC clone distribution information can be
 found through the I.M.A.G.E. Consortium/LLNL at:
<http://image.llnl.gov>
 Plate: LHAM10394 row: b column: 22
 High quality sequence stop: 698.
 Location/Qualifiers
 1..803

FEATURES

source

/organism="Mus musculus"
 /db_xref="taxon:10090"
 /clone="IMAGE:451157"
 /clone_lib="NIH-MGC_94"
 /tissue_type="retina"
 /lab_host="DH10B (Phage-resistant)"
 /note="Organ: eye; Vector: pCMV-Sport6; Site_1: NotI;
 Site_2: SalI; Cloned unidirectionally; Oligo-dT primed.
 Average insert size 3.3 kb. Library enriched for
 full-length clones and constructed by Life Technologies.
 Note: this is a NIH-MGC library."
 BASE COUNT 240 a 162 c 205 g 196 t

Query Match 11.9%; Score 509.6; DB 12; Length 803;
 Best Local Similarity 80.8%; Pred. No. 3.2e-78;
 Matches 643; Conservative 0; Mismatches 149; Indels 4; Gaps 4;

QY 729 TATCTTCATTTAGTATTAAGAACTCTTGCTATGCAAAAGCTGGAGCACTAGCTAG 788
 DB 10 TGTCTTCATTTAGTATTAAGAACTCTTGCTATGCAAAAGCTGGAGCACTAGCTAG 69
 QY 789 AAGTCTTACACATTCAGAACTGTGATTTGCTTTGAGAGCAAAAGAAAGAACTTGA 848

||||| 128
Db AACTCTAGAGCCATCAAGACTGATG -GTTGGAGGCAAAAGAGAACTTGA 128
QY 849 GGTACAAACAAATTTAGAAAGCAATTTGGATTAAGAACTATACAGGCCA 908
Db 129 GGTACAAATACAACTTGAGAAAGCTTAAGAGCTGGGATTAAGAACTATACAGGCCA 188
QY 909 ACATTTCTATGTTGGCGCTTTCTTATGATCTATGATCATATGCTCTGGCTTCTG 968
Db 189 ACATCTCATGGGAGGAGCTTTCTCTTATCTATGATCATATGCTCTGGATCTG 248
QY 969 ATGGAGCTCTGTTGCTCTCTCCAGTAATTTCTATGAGACAGTACTACTCTCT 1028
Db 249 ATGGAGCTCTCTGTTGCTCTCTCCAGTAATTTCTATGAGACAGTACTACTCTCT 307
QY 1029 TTTCTGATTAATTTGGGCTTTTATGATTTGAGACAGGCAATCCAGCATTTG 1088
Db 308 TTTCCGTTTATTTGAGACATTCAGTTGAGACAGGCAATCCAGCATTTGAG 367
QY 1089 CAACGCAAGAGAGAGCTTATGAAATCTTCAAGATTAATGACATTAACCAAGCAT 1148
Db 368 CCAATGCGAGGAGAGAGCTTATGAAATCTTCAAGATTAATGACATTAACCAAGCAT 427
QY 1149 ACAGCTATTCAGAGTGGACATTAACCAAGATTAATGAGGAATTTGGAATTTG 1208
Db 428 ACAGCTTCTCAAGAGTGGACATTAACCAAGATTAATGAGGAATTTGGAATTTG 487
QY 1209 ATGTTCACTTACATTCCTTCTGAAAGAGTAAATGATTTAAAGGCTCAACCTG 1268
Db 488 ATATTCACCTTACATTCCTTCTGAAAGAGTAAATGATTTAAAGGCTCAACCTG 547
QY 1269 AGTTTCAGTGGGAGAGAGTGGGCTGTTGGGAACAGTGGCTGGGGAAGAGCAG 1328
Db 548 AGTTTCAGTGGGAGAGAGTGGGCTGTTGGGAACAGTGGCTGGGGAAGAGCAG 607
QY 1329 CCGTGCAGTGTATGAGAGAGTGTATGACCCAGAGATGGATGCTGATTTGATG 1388
Db 608 CTGTCCAGCTGATGCAAGAGCTGTACAGCCC -TGAATGGATGGTGCATTCAGAG 666
QY 1389 AGGACATTTAGGACAT -AAATGTAAGGCAATTCGGAATTTACTGTTGTGATG 1447
Db 667 AGGACATTTAGGACATTTAGGATTTGAGAGGAGATTCAGGAGTGTGATGATG 726
QY 1448 GAGCTGTGTTGTTGCCACAGATAGCTGAATTCCTGATGCGCGGAATATGTC 1507
Db 727 GAACCTGCTGTTGTTGCCACAGATAGCTGCGGAATATTCCTGATGCGCGGAATATG 786
QY 1508 ACCATGGATGATG 1523
Db 787 GTTCTGTTGTTGATG 802

RESULT 13
BM053000 587 bp mRNA linear EST 12-MAR-2002
LOCUS 169807.y3 Melton Normalized Human Islet 4 N4-HIS 1 Homo sapiens
DEFINITION cDNA clone IMAGE:5671980 5' similar to SW:MDRI.HUMAN P08183
MOLPROP RESISTANCE PROTEIN 1 ; mRNA sequence.
ACCESSION BM053000
VERSION BM053000.1 GI:1680894
KEYWORDS EST.
SOURCE human.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.
REFERENCE 1 (bases 1 to 587)
AUTHORS Melton,D., Brown,J., Keny,G., Permutt,A., Lee,C., Kaestner,K.,
Lemishka,I., Seearce,M., Brestelli,J., Gradwohl,G., Clifton,S.,
Hillier,L., Marra,M., Pape,D., Wylie,T., Martin,J., Blaisdin,A.,
Schmitt,L., Theising,B., Ritter,E., Ronko,I., Bennett,J., Cardenas
'M., Gibbons,M., McCann,R., Cole,R., Tsagarashwill,R., Williams,T.,
Jackson,X. and Bowers,Y.
Endocrine Pancreas Consortium

JOURNAL Unpublished (2000)
COMMENT Contact: Douglas Melton, Klaus H. Kaestner, & Hiroshi Inoue
Endocrine Pancreas Consortium
Harvard University, Howard Hughes Medical Institute
Dept of Molecular and Cellular Biology, 7 Divinity Ave, Cambridge,
MA 02138
Tel: 617-495-1812
Fax: 617-495-8557
Email: dmelton@bioh.harvard.edu
Library was constructed by Dr. Douglas Melton DNA sequencing by:
Washington University Genome Sequencing Center for information on
obtaining a clone please contact: Juliana Brown
(brownj@fas.harvard.edu) This sequence now available from the IMAGE
consortium for clone orders contact: info@image.llnl.gov
High quality sequence stop: 437.

FEATURES
source
1..587
/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone="IMAGE:5671980"
/clone_lib="Melton Normalized Human Islet 4 N4-HIS 1"
/sex="Both"
/tissue_type="Islets of Langerhans"
/dev_stage="Adult"
/lab_host="DH10B"
/note="Organ: Pancreas; Vector: pSPORT1; Site:1; Not 1;
Site:2; Sal 1; Starting library constructed using
Superscript Plasmid Library kit (Life Technologies). cDNA
made by oligo-dT priming. Size-selected by column
fractionation; average insert size 1.08 kb. Library was
amplified once on solid support and plasmid DNA from
library was prepared. The library DNA was normalized by
method #4 from Bonadio, Lennon, and Soares 1996 Genome
Research 6:791-806; 0.5 microgram single-stranded library
plasmid DNA was mixed with 5 micrograms PCR product
representing library inserts and hybridized to an Acot of
20 Single-stranded (unhybridized) plasmids were isolated
by hydroxyapatite chromatography and used to make this
library"

BASE COUNT 142 a 138 c 161 g 145 t 1 others
ORIGIN
Query Match 11.2%; Score 479; DB 13; Length 587;
Best Local Similarity 89.4%; Pred. No. 7.1e-73;
Matches 515; Conservative 0; Mismatches 61; Indels 0; Gaps 0;
QY 2862 CAATGATTAATTTTCTTCTATGCTGCTCCGTTGCTGCTACTGCTGCAATG 2921
Db 1 CAATGATTAATTTTCTTCTATGCTGCTCCGTTGCTGCTACTGCTGCAATG 60
QY 2922 AGTTCAATGACCTTTCAGATGTTCTTTTGGTATTTCTCAGCTATTTGCTTGGTCCATG 2981
Db 61 AACTCATGAGCTTTGAGATGTTCTGTTAGTATTTTCAGCTGTTGCTTGGTCCATG 120
QY 2982 CAGTGGGGAGAGTCACTTATTTGCTCCCTGATGATGCCAAAGCCAAATATCAGAGCC 3041
Db 121 CCGGAGGAGAGTCACTTATTTGCTCCCTGATGATGCCAAAGCCAAATATCAGAGCC 180
QY 3042 ACCTCATGATGATCAATGAAAGAGCCCTGATTTGACAGTACAGCCCTCAGCCCTCA 3101
Db 181 ACATCATGATGATCAATGAAAGAGCCCTTGTATGATGACAGTACAGCCGGAAGCCCTCA 240
QY 3102 AGCCAAATACCTTGAAGAGAAATGATGATTTAATGAGTGGTTCATCAATCTCCACTG 3161
Db 241 TGCCGACATTTGGAAGAAATGATGATTTGATGATTTGATTTCAATCTCCACTG 300
QY 3162 GACCAAGATCCCGCTCCAGAGGCTGAGCTTCAGAGTGAAGAGGCCAGAGCTTG 3221
Db 301 GACCGGATCCAGTCTCTTCAGAGGAGTGAAGAGGCCAGAGCTTG 360
QY 3222 CCTCTGATGATGAGTGGCTGTGGAGAGCAGTGTGTCAGTCCAGAGGCTTCT 3281
Db 361 CTCTGATGAGTGGCTGTGGAGAGCAGAGTGTGTCAGTCCAGAGGCTTCT 420

QY 3282 ATGACCCCTGGCTGCTTAATGATGACCAAGATAAGACACCTGATGTC 3341
 DB 421 ACGACCCCTGGAGGGAAGTGTGCTGATGCAAAATAAGCCAGCAATGTC 480
 QY 3342 AGTGGCTCCGAGCACACCTGGGCACTGCTGTCTCAGAGACCCACTCTGTTGACTGACGA 3401
 DB 481 AGTGGCTCCGAGCACACCTGGGCACTGCTGTCTCAGAGACCCACTCTGTTGACTGACGA 540
 QY 3402 TTGGCGAAGACATGCTTATGAGACACGCGGG 3437
 DB 541 TTGCTGAGAACATTCCTATGAGACACGCGGG 576

RESULT 14
 AV616675 605 bp mRNA linear EST 28-NOV-2001
 LOCUS AV616675 Bos taurus ovary fetus Bos taurus cDNA clone E10V006G02
 DEFINITION 5', mRNA sequence.
 ACCESSION AV616675
 VERSION AV616675.1 GI:9752345
 KEYWORDS EST.
 SOURCE cow.
 ORGANISM Bos taurus
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Butheria; Cetartiodactyla; Ruminantia; Pecora; Bovidae;
 Bovidae; Bovinae; Bos.

REFERENCE 1 (bases 1 to 605)
 Takasuga,A., Hirotsune,S., Itoh,R., Jitohzono,A., Suzuki,H., Aso,H.
 and Sugimoto,Y.
 Establishment of a high throughput EST sequencing system using
 poly(A) tail-removed cDNA libraries and determination of 36,000
 bovine ESTs

TITLE Nucleic Acids Res. 29 (22), E108 (2001)
 JOURNAL 21570554
 MEDLINE Contact: Yoshikazu Sugimoto
 COMMENT Animal Genetics Division
 Shitakawa Institute of Animal Genetics
 Odeokura, Nishigo, Nishi-shitakawa, Fukushima 961-8061, Japan
 Tel: 81-248-25-5641
 Fax: 81-248-25-5725

Email: kazusugie@cocoa.ocn.ne.jp
 Single pass sequencing.
 This clone was obtained from a polyA-deleted cDNA library.

FEATURES
 source
 1..605
 Location/Qualifiers
 /organism="Bos taurus"
 /db_xref="taxon:9913"
 /clone_id="E10V006G02"
 /clone_lib="Bos taurus ovary fetus"
 /tissue_type="ovary"
 /dev_stage="fetus"
 /lab_host="DH10B"
 /note="Vector: pZL1; Site_1: SalI; Site_2: NotI; Poly A
 was deleted from a NotI site"

BASE COUNT 158 a 128 c 148 g 171 t
 ORIGIN

Query Match 11.2%; Score 479; DB 10; Length 605;
 Best Local Similarity 88.9%; Pred. No. 7e-73;
 Matches 518; Conservative 0; Mismatches 65; Indels 0; Gaps 0;

QY 2556 CAGGCAATTAATACCTTAATCATGTTGGCAATTAACACTTTTACCTAGCAATTTG 2615
 DB 10 CGGCAATTAATATCCATCATCTATGCTGCGACAGTTAACTTCCTCTGTTAGCAATTTG 69
 QY 2616 TACCCATCATTCGATGAGAGAGAGTGTGAATGAAATGTTGTCTGAGACAGCACTGA 2675
 DB 70 TACCCATCATTCGATGAGAGAGAGTGTGAATGAAATGTTGTCTGAGACAGCACTGA 129
 QY 2676 AAGATAGAAAAGAGCTAGAGAGAGCTGGAGATGTTAGACAGAGCCATGCAAAACTTCC 2735
 DB 130 AAGATAGAAAAGAGCTAGAGAGAGCTGGAGATGTTAGACAGAGCCATGCAAAACTTCC 189

QY 2736 GAAGTGTGTTCTTTGACATCGGAGAGAGAGTTGAATACATGATGACAGAGTTGTC 2795
 DB 190 GGACTGTGTTCTTTTACTCGGGAGAGAGAGTTGAATATATGATGATGCCAGAGTTGTC 249
 QY 2796 AAGTACCATACAGAAACTCTTTGAGGAAAGACACACATCTTGGGGCTCATTTTGCATCA 2855
 DB 250 AGTACCATACAGAAACTCTTTGAGGAAAGACACATCTTGGAAATTAATTTGCCCTTCA 309
 QY 2856 CCCAGCAATGATGATATTTTCTATGCTGCTGCTTCCGGTTGGTGCCCTACTTGTGTG 2915
 DB 310 CCCAGCAATGATGATATTTCTCTATGCTGCTGCTTCCGGTTGGTGCCCTACTTGTGTG 369
 QY 2916 CAAATGAGTTCATGAACATTCATGATGTTGATCTGATCTCAGCATATGCTTGTGTC 2975
 DB 370 CACAAAGCATCATGAGAGTTTCAAGATGTTCTCTTACTGTTCTCAGCATATGCTTGTGTC 429
 QY 2976 CCATGGCAGTGGGGCGAGTCACTTTCCTGCTGACTATGCCAAAGCCAAAGATGAG 3035
 DB 430 CCATGGCTGTGGGACAGTCACTTTCCTGCTGACTATGCCAAAGCCAAAGATGAG 489
 QY 3036 CAGCCACGTCATCATGATCATTTGAAAAAGCCCTCTGATTCAGAGCTACAGCCCTACG 3095
 DB 490 CAGCCACGTCATCATATTCATTTGAAAAATTCCTCTGATTCAGAGCTACAGAGAG 549
 QY 3096 GCCTCAGCCAAATACGTTGGAAGGAAATGTGACATTTAATGA 3138
 DB 550 GCCTTAAGCCGAGTACAGTGAAGGAAACGTGGCATTTAATGA 592

RESULT 15
 BF313560 795 bp mRNA linear EST 21-NOV-2000
 LOCUS BF313560
 DEFINITION 601900192P1 NIH_MGC_19 Homo sapiens cDNA clone IMAGE:4129222 5',
 mRNA sequence.
 ACCESSION BF313560
 VERSION BF313560.1 GI:11261583
 KEYWORDS EST.
 SOURCE human.
 ORGANISM Homo sapiens
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Primates; Catarrhini; Hominiidae; Homo.

REFERENCE 1 (bases 1 to 795)
 NIH-MGC http://mgc.nci.nih.gov/.
 TITLES National Institutes of Health, Mammalian Gene Collection (MGC)
 JOURNAL Unpublished (1999)
 COMMENT Contact: Robert Strausberg, Ph.D.
 Email: c9apbs-r@mail.nih.gov
 Tissue Procurement: ATCC
 CDNA Library Preparation: Ling Hong/Rubin Laboratory
 CDNA Library Arrayed by: The I.M.A.G.E. Consortium (LNLN)
 DNA Sequencing by: Incyte Genomics, Inc.
 Clone distribution: MGC clone distribution information can be
 found through the I.M.A.G.E. Consortium/LNLN at: Image.lnl.gov
 Plate: LNCM1023 row: h column: 23
 High quality sequence stop: 674.

FEATURES
 source
 1..795
 Location/Qualifiers
 /organism="Homo sapiens"
 /db_xref="taxon:9606"
 /clone_id="IMAGE:4129222"
 /clone_lib="NIH_MGC_19"
 /tissue_type="neuroblastoma"
 /lab_host="DH10B (phage-resistant)"
 /note="Organ: brain; Vector: pORF; Site_1: XhoI; Site_2:
 EcoRI; CDNA made by oligo-dT priming. Directionally
 cloned into EcoRI/XhoI sites using the following 5'
 adaptor: GGCACGAG(G). Library constructed by Ling Hong
 in the laboratory of Gerald M. Rubin (University of
 California, Berkeley) using ZAP-CDNA (Life Technologies)
 (Stratagene) and Superscript II RT (Life Technologies).
 Note: this is a NIH_MGC library."

BASE COUNT 213 a 189 c 236 g 157 t

ORIGIN

Query Match 11.1%; Score 475.2; DB 12; Length 795;
Best Local Similarity 83.6%; Pred. No. 2.8e-72;
Matches 612; Conservative 0; Mismatches 108; Indels 12; Gaps 6;

QY 3065 AGCCCTCTGATTGACACCTCTACAGCCCTCAGCGCCCTCAAGCCCAATTCGTTGGAGGAAAT 3124
DB 2 AACCCCTTTGATGACACCTACAGCAGCGAAGGCGCTTAATGCCAGAC-ATGGAGGAAAT 60
QY 3125 GTGACATTTAATGAGTGTGTTCAACTATCCACTCGACACAGACATCCCCGCTCCAG 3184
DB 61 GTCACTATGGGAAG--TGATTTCACTATCCACCAGCCGACATCCAGTCTTCAG 118
QY 3185 GGGCTGACCTCGAGGTGAGAGAGAGGCGCACACCTGGCCCTGCTAGGTAGCACTGCTGT 3244
DB 119 GGACTGACCTGAGGTGAGAGAGAGGCGCACAGCTGGCTGTGGTGGCAGACAGTGGCTGT 178
QY 3245 GGGAGAGACACAGTGTTCAGCTCTGAGCGCTTCTATGACCCCTTGCTGGTTCAGTG 3304
DB 179 GGGAGAGACACAGTGTTCAGCTCTGAGCGGTTCAGACCCCTTGGCAGGAAAGTG 238
QY 3305 CTATTTGATGGCAAGAGATTAAGCACCTGATATGCTCAGTGGCTCCGAGCACACTGGGC 3364
DB 239 CTGCTTGAATGGCAAGAAATTAAGCACTGATATGCTCAGTGGCTCCGAGCACACTGGGC 298
QY 3365 ATCGTGTCTCAGAGACCCCTCTGTTGCTGACAGATTCGAGCAATTCGCTATGGA 3424
DB 299 ATCGTGTCTCAGAGACCCCTCTGTTGCTGACAGATTCGCTGAGAACTTGCCTATGGA 358
QY 3425 GACAACAGCGCGGTGTATCATGATGAGAGATTATGACAGCAGCCAAAGAGGCCACATA 3484
DB 359 GACAACAGCGCGGTGTATCATGATGAGAGATTATGACAGCAGCCAAAGAGGCCACATA 418
QY 3485 CACCCTTATGAGACACTCCCTGAGAAATPACACACAGATGAGAGACAAAGAAC 3544
DB 419 CAGCCTTATGAGACACTCCCTGAGAAATPACACACTAAAGTAGAGAACAAAGAACT 478
QY 3545 CAGCTCTGTGTGGCCAGAAACAGCGCATTTGCCATAGCTCGGCTCTTGTAGACAGCT 3604
DB 479 CAGCTCTGTGTGGCCAGAAACAGCGCATTTGCCATAGCTCGGCTCTTGTAGACAGCT 538
QY 3605 CATATTTTGTCTTTGGATGAGCTACATGAGCTGAGATACAGAAAGTGAAGGT-TGT 3663
DB 539 CATATTTT--GCTTTGATGAGAGCCACGCTCAGCTGAGATACAGAAAGTGAAGGTGCT 596
QY 3664 CCAAGAGAGCCCTGAGACAAAGCCAGAGAGGCGCACCTGATGTGATCGCCACCGCT 3723
DB 597 CCAAGAGAGCCCTGAGACAAAGCCAGAGAGGCGCACCTGATGTGATGTGCACCGC---T 652
QY 3724 GTCCACATCCAGATGAGATTATAGTGGTTCAGAAATGGCAAGTCAAGAGCA 3783
DB 653 GTCCACATCCAGATGAGACTT-ATAGTGTGTTTCAGAAATGGCAGAGTCCAGGGGCG 710
QY 3784 TGGCAGACATCA 3795
DB 711 ATGGACGATCA 722

Search completed: December 10, 2002, 05:39:28
Job time : 29569 secs

OM protein - protein search, using sw model

(without alignments)
184.136 Million cell updates/sec

Sequence: 1 MDPEGGRKGSAEKNFWKMGK.....LLAQGIYFSMVSYQAGAKR 1281

Gapop 10.0 , Gapext 0.5

Total number of hits satisfying chosen parameters: 908470

Post-processing: Minimum Match 08

```

Post-processing:  Minimum Match 0%
                  Maximum Match 100%
                  Listing first 45 summaries

```

Database

23.	/SIDS1.gcgdata.genseq.genseq_emb1/AA1980.DAT
22.	/SIDS1.gcgdata.genseq.genseq_emb1/AA1981.DAT
21.	/SIDS1.gcgdata.genseq.genseq_emb1/AA1982.DAT
20.	/SIDS1.gcgdata.genseq.genseq_emb1/AA1983.DAT
19.	/SIDS1.gcgdata.genseq.genseq_emb1/AA1984.DAT
18.	/SIDS1.gcgdata.genseq.genseq_emb1/AA1985.DAT
17.	/SIDS1.gcgdata.genseq.genseq_emb1/AA1986.DAT
16.	/SIDS1.gcgdata.genseq.genseq_emb1/AA1987.DAT
15.	/SIDS1.gcgdata.genseq.genseq_emb1/AA1988.DAT
14.	/SIDS1.gcgdata.genseq.genseq_emb1/AA1989.DAT
13.	/SIDS1.gcgdata.genseq.genseq_emb1/AA1990.DAT
12.	/SIDS1.gcgdata.genseq.genseq_emb1/AA1991.DAT
11.	/SIDS1.gcgdata.genseq.genseq_emb1/AA1992.DAT
10.	/SIDS1.gcgdata.genseq.genseq_emb1/AA1993.DAT
9.	/SIDS1.gcgdata.genseq.genseq_emb1/AA1994.DAT
8.	/SIDS1.gcgdata.genseq.genseq_emb1/AA1995.DAT
7.	/SIDS1.gcgdata.genseq.genseq_emb1/AA1996.DAT
6.	/SIDS1.gcgdata.genseq.genseq_emb1/AA1997.DAT
5.	/SIDS1.gcgdata.genseq.genseq_emb1/AA1998.DAT
4.	/SIDS1.gcgdata.genseq.genseq_emb1/AA1999.DAT
3.	/SIDS1.gcgdata.genseq.genseq_emb1/AA2000.DAT
2.	/SIDS1.gcgdata.genseq.genseq_emb1/AA2001.DAT
1.	/SIDS1.gcgdata.genseq.genseq_emb1/AA2002.DAT

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB	ID	Description
1	6477	100.0	1281	22	AAE00303	Dog (Pgp) P-glycoprotein
2	6469	99.9	1281	22	AAE00308	Dog P-glycoprotein
3	6463	99.8	1281	22	AAE00309	Dog P-glycoprotein
4	6456	99.7	1281	22	AAE00310	Dog P-glycoprotein
5	6420.5	99.1	1280	22	AAE00310	Dog P-glycoprotein
6	6420.5	99.1	1280	22	AAE00304	Dog P-glycoprotein
7	5861	90.5	1280	22	AAE00307	Dog P-glycoprotein
8	5861	90.5	1280	21	AAE00303	Dog P-glycoprotein
9	5861	90.5	1280	22	AAE00308	Dog P-glycoprotein
10	5861	90.5	1280	23	AAE00304	Dog P-glycoprotein

11	5861	90.5	1280	23	ABBO7266	Human BCRP protein
12	5859	90.5	1280	23	AAE18967	Human P-gp mutant,
13	5857	90.4	1280	22	AAE18967	Cynomolgous monkey
14	5856	90.4	1280	23	AAE18968	Human P-gp mutant,
15	5854.5	90.4	1283	22	AAE18965	Cynomolgous monkey
16	5854	90.4	1280	23	AAE18966	Human P-gp mutant,
17	5853	90.4	1280	23	AAE18969	Human P-gp mutant,
18	5853	90.4	1280	23	AAE18970	Human P-gp mutant,
19	5853	90.4	1280	23	AAE18973	Human P-gp mutant,
20	5852	90.4	1280	21	AAE18917	Human P-gp mutant,
22	5852	90.4	1280	22	AAE18966	Human P-glycoprotein
22	5852	90.4	1280	22	AAE18966	Human P-glycoprotein
24	5852	90.4	1280	23	AAE24211	Human P-glycoprotein
25	5852	90.4	1280	23	AAE18965	Human MDR protein,
26	5850	90.3	1280	23	ABBO7267	Human P-gp mutant,
27	5849	90.3	1280	23	AAE18977	Human BCRP protein
28	5849	90.3	1280	8	AAE18977	Human P-gp mutant,
29	5847	90.3	1280	23	AAE18971	Sequence encoded b
30	5846	90.3	1280	23	AAE18978	Human P-gp mutant,
31	5845	90.2	1280	11	AAE18968	Human P-gp mutant,
32	5844.5	90.2	1279	22	AAE18997	Protein encoded by
33	5844.5	90.2	1279	22	AAE00307	Human P-gp mutant,
34	5844	90.2	1280	23	AAE18974	Human P-glycoprotein
35	5844	90.2	1280	23	AAE18979	Human P-glycoprotein
36	5844	90.2	1280	23	AAE18960	Human P-gp mutant,
37	5841	90.2	1280	23	AAE18972	Human P-gp mutant,
38	5840	90.2	1280	23	AAE18972	Human P-gp mutant,
39	5839	90.1	1280	15	AAE18975	Human P-gp mutant,
40	5836	90.1	1280	14	AAE44397	Human P-glycoprotein
41	5836	90.1	1280	22	AAE00347	Sequence encoded b
42	5835.5	90.1	1279	19	AAE00394	Human multidrug re
44	5832	90.0	1280	23	AAE18976	Mutated human P-gl
44	5832	90.0	1280	23	AAE18976	Human P-gp mutant,
45	5825	89.9	1280	19	AAE48997	Wild-type human P
				23	AAE07493	Human P-glycoprotein

ALIGNMENTS

RESULT 1	
AAE00303	
ID	AAE00303 standard; Protein; 1281 AA
VY	

AC AAEE00303;

DT 13-JUN-2001 (first entry)
XY

Dog (PGP) P-glycoprotein (genotype C) #1

XX drug bioavailability; transgenic animal; genetic model

OS Canis familiaris.
YY

PN WO200123540-A2.
XX

PD 05-APR-2001

PF 28-SEP-2000; 2000WO-US26767
YY

PR 28-SEP-1999; 99US-0156510

PA (GENT-) GENTEST CORP.
XX

Stocker PJ, Steimel-crespi DT, Crespi CL, Reif TC, Patten CJ;

DR WPI; 2001-235373/24.

DR N-PSDB; AAD03488.

PT New dog P-glycoproteins (PGP) and their encoding nucleic acids, useful

PT PGP inhibitors -

XX Claim 17; Page 64-66; 11pp; English.
CC The invention relates to dog P-glycoprotein (PGP) also referred
CC as multidrug transporter (MDR1) and nucleic acids encoding them.
CC The invention also includes fragments and biologically functional
CC variants of dog P-glycoprotein. PGP and their nucleic acids are
CC useful for determining the bioavailability of drugs and for
CC screening PGP inhibitors. They are useful for the diagnosis and
CC treatment of conditions characterized by PGP activity, by
CC reducing or increasing PGP activity in a cell. PGP nucleic acids
CC are used as oligonucleotide probes. Complements of PGP nucleic
CC acids are useful as antisense oligonucleotides, to induce a PGP
CC 'knockout' phenotype. They are used to prepare a non-human
CC transgenic animal, which are valuable as genetic models for
CC human diseases.
CC The present sequence is dog P-glycoprotein (PGP) also referred
CC as genotype C protein. The PGP enzyme functions as an efflux
CC pump exporting small molecules across the cell membrane. This
CC enzyme is a member of the ABC transporter family.

XX Sequence 1281 AA:

Query Match 100.0%; Score 6477; DB 22; Length 1281;
Best Local Similarity 100.0%; Pred. No. 0; Mismatches 0; Indels 0; Gaps 0;
Matches 1281; Conservative 0;

QY 1 MDPEGGRKSAEKNFMWKKSKKNEKKKPPVSTFAMERYSNMLDRLVMVTMAII 60
DB 1 MDPEGGRKSAEKNFMWKKSKKNEKKKPPVSTFAMERYSNMLDRLVMVTMAII 60
QY 61 HGAALPLMLLVFGNMTDSFANAGISRNKTPPVYINSEINNNQHFNHLEEMTYAAY 120
DB 61 HGAALPLMLLVFGNMTDSFANAGISRNKTPPVYINSEINNNQHFNHLEEMTYAAY 120
QY 121 SGISAGVLAAYIOVSFWCLAGROILIKRQPFHAIMROEIGMFDVHDVDELNTLTD 180
DB 121 SGISAGVLAAYIOVSFWCLAGROILIKRQPFHAIMROEIGMFDVHDVDELNTLTD 180
QY 181 VSKINEGIGDKIGMFHSHIATFTGTIVGTFRGKMLTVLILAIAPVLGISAIAWAKILSS 240
DB 181 VSKINEGIGDKIGMFHSHIATFTGTIVGTFRGKMLTVLILAIAPVLGISAIAWAKILSS 240
QY 241 FTDKELAYAKAGAAVEEVLAAITRTVAFPGOKKELERKYNKEAGKIGIKKAITANIS 300
DB 241 FTDKELAYAKAGAAVEEVLAAITRTVAFPGOKKELERKYNKEAGKIGIKKAITANIS 300
QY 301 IGAAPLTIYASYALAFWGTSLVLSSEYSIGOVLTVEFSVLIGAFSIGASPSIEAFANA 360
DB 301 IGAAPLTIYASYALAFWGTSLVLSSEYSIGOVLTVEFSVLIGAFSIGASPSIEAFANA 360
QY 361 RGAAYEIKTIIDNKPISIDYSKSHKPDNKGNEFKNVHPSYPSRREYVILGLMLKYQ 420
DB 361 RGAAYEIKTIIDNKPISIDYSKSHKPDNKGNEFKNVHPSYPSRREYVILGLMLKYQ 420
QY 421 SGOYVALVNGSCGKSTVOLMORLYDPTDGMVCIDGQDIRTINVRILREITGVSOEY 480
DB 421 SGOYVALVNGSCGKSTVOLMORLYDPTDGMVCIDGQDIRTINVRILREITGVSOEY 480
QY 481 LEATTIAENIRYGRNVTMDIEKAVKANAYDFIMKLPNKFTLVBERAOISGQKOR 540
DB 481 LEATTIAENIRYGRNVTMDIEKAVKANAYDFIMKLPNKFTLVBERAOISGQKOR 540
QY 541 IAAIALVLRNPKIILLDDATSAIDTESPAVVOVALDKARKGRTIVIAHRLSTVYRNDVI 600
DB 541 IAAIALVLRNPKIILLDDATSAIDTESPAVVOVALDKARKGRTIVIAHRLSTVYRNDVI 600
QY 601 AGFDGVIYEKGNHDELKKEGIYFKLVTMOTRGNIELEENATGESKSESDALEMSPKDS 660
DB 601 AGFDGVIYEKGNHDELKKEGIYFKLVTMOTRGNIELEENATGESKSESDALEMSPKDS 660
QY 661 GSSLIKRSTRSRSHAPGQDRKLTCKEDLNNVPPVSFKILKLNSTEMPYFVVGIFCA 720
DB 661 GSSLIKRSTRSRSHAPGQDRKLTCKEDLNNVPPVSFKILKLNSTEMPYFVVGIFCA 720

DB 661 GSSLIKRSTRSRSHAPGQDRKLTCKEDLNNVPPVSFKILKLNSTEMPYFVVGIFCA 720
QY 721 IINGLIQPAFSTIIBRIIGITFRDEDPETKRONNMNSVFLVIGIISFTFFLOGFTFG 780
DB 721 IINGLIQPAFSTIIBRIIGITFRDEDPETKRONNMNSVFLVIGIISFTFFLOGFTFG 780
QY 781 KAGELITKRLRMVRRSMRLRODSWFDPRKMTGALTTRLANDAAQVKAIGSLAVITQ 840
DB 781 KAGELITKRLRMVRRSMRLRODSWFDPRKMTGALTTRLANDAAQVKAIGSLAVITQ 840
QY 841 NIANLGTGIIISILYGMOTLLILAIIVPITAIAGYVEKMLSGALDKKELBAGKIAT 900
DB 841 NIANLGTGIIISILYGMOTLLILAIIVPITAIAGYVEKMLSGALDKKELBAGKIAT 900
QY 901 EATENRTVASTLRQKFEYMAQSLQVPRNSLRKAHIFGVSPSITQAMMYFYACFR 960
DB 901 EATENRTVASTLRQKFEYMAQSLQVPRNSLRKAHIFGVSPSITQAMMYFYACFR 960
QY 961 FGAYLVANEFMNFODVLVFSALVFGAMAVGVSSFADYAKAKVSAHVMIIEKSPLI 1020
DB 961 FGAYLVANEFMNFODVLVFSALVFGAMAVGVSSFADYAKAKVSAHVMIIEKSPLI 1020
QY 1021 DSYSPHGLKPNLTIEGVTNEVEVFNPPRPDIIPVLOGLSLEVKKGOTLALVGSSGCKST 1080
DB 1021 DSYSPHGLKPNLTIEGVTNEVEVFNPPRPDIIPVLOGLSLEVKKGOTLALVGSSGCKST 1080
QY 1081 VVOLLERFYDPLAGSVLIDGKEIKHLNVOMLRANLGIYSQEPILDPDCAIENIAYGNSR 1140
DB 1081 VVOLLERFYDPLAGSVLIDGKEIKHLNVOMLRANLGIYSQEPILDPDCAIENIAYGNSR 1140
QY 1141 VVSHHEIMOAKKANIHFIETLPEKYNTRVGDKGQOLSGCGKOKIATARALVROPHTL 1200
DB 1141 VVSHHEIMOAKKANIHFIETLPEKYNTRVGDKGQOLSGCGKOKIATARALVROPHTL 1200
QY 1201 LDEATSAIDTESKRYVOEALDKAREGRTCIYIAHRLSTIONADLLIVPONGKVKRSHGTH 1260
DB 1201 LDEATSAIDTESKRYVOEALDKAREGRTCIYIAHRLSTIONADLLIVPONGKVKRSHGTH 1260
QY 1261 QLLAQKGIYFSMVSVOGAKR 1281
DB 1261 QLLAQKGIYFSMVSVOGAKR 1281

RESULT 2
AAE00308
ID AAE00308 standard; Protein; 1281 AA.
XX
AC AAE00308;
DT 13-JUN-2001 (first entry)
XX
DE Dog P-glycoprotein (PGP) allelic variant (Genotype A).
XX
KW Dog; P-glycoprotein allelic variant; PGP; multidrug transporter;
KW MDR1; drug bioavailability; transgenic animal; genetic model.
XX
OS Canis familiaris.
XX
FH Key location/Qualifiers
FH Misc-difference 197 /note= "His of genotypc substituted by Gln"
FT
PI WO200123540-A2.
PN
PD 05-APR-2001.
XX
PE 28-SEP-2000; 2000WO-US26767.
XX
PR 28-SEP-1999; 99US-0156510.
XX
PA (GENT-) GENTEST CORP.
XX
PI Stocker PJ, Steimel-crespi DT, Crespi CL, Relif TC, Patten CJ;

The invention relates to dog P-glycoprotein (PGP) also referred as multidrug transporter (MDR1) and nucleic acids encoding them. The invention also includes fragments and biologically functional variants of dog P-glycoprotein. PGP and their nucleic acids are useful for determining the bioavailability of drugs and for screening PGP inhibitors. They are useful for the diagnosis and treatment of conditions characterised by PGP activity, by reducing or increasing PGP activity in a cell. PGP nucleic acids are used as oligonucleotide probes. Complements of PGP nucleic acids are useful as antisense oligonucleotides, to induce a PGP 'knockout' phenotype. They are used to prepare a non-human transgenic animal, which are valuable as genetic models for human diseases.

The present sequence is dog P-glycoprotein (PGP) allelic variant. This sequence is also referred as Genotype A protein. The PGP enzyme functions as an efflux pump exporting small molecules across the cell membrane. This enzyme is a member of the ABC transporter family.

Sequence 1281 AA:

0.

Db	541	IAIARALYRNPKILL	LD	DEAT	SALD	SEAV	VOVALD	KARKGRTT	IVIAHRLST	VRNADVI	600																																												
Oy	601	AGFD	DGV	IV	EKG	NHDELM	EKGI	YF	FLV	MTGR	NGNEIELENA	TGSESD	AL	EM	PKDS	660																																							
Db	601	AGFD	DGV	IV	EKG	NHDELM	EKGI	YF	FLV	MTGR	NGNEIELENA	TGSESD	AL	EM	PKDS	660																																							
Oy	661	GSS	LTKRR	SR	RR	SI	HA	P	Q	D	R	K	L	T	K	E	D	LN	EN	VP	VS	FW	RL	KL	NST	EM	P	Y	FW	IF	CA	720																							
Db	661	GSS	LTKRR	SR	RR	SI	HA	P	Q	D	R	K	L	T	K	E	D	LN	EN	VP	VS	FW	RL	KL	NST	EM	P	Y	FW	IF	CA	720																							
Oy	721	I	NGG	L	P	A	S	I	I	F	S	R	L	I	G	T	F	R	D	E	P	E	P	T	K	R	N	S	M	S	V	L	V	L	G	I	S	F	T	F	L	O	G	T	T	E	G	780							
Db	721	I	NGG	L	P	A	S	I	I	F	S	R	L	I	G	T	F	R	D	E	P	E	P	T	K	R	N	S	M	S	V	L	V	L	G	I	S	F	T	F	L	O	G	T	T	E	G	780							
Oy	781	K	A	E	I	L	T	K	L	R	L	R	V	E	R	S	M	L	R	O	V	S	F	E	D	D	P	K	N	T	G	A	L	T	R	L	N	D	A	O	A	G	K	A	I	G	S	R	L	A	V	I	T	Q	840
Db	781	K	A	E	I	L	T	K	L	R	L	R	V	E	R	S	M	L	R	O	V	S	F	E	D	D	P	K	N	T	G	A	L	T	R	L	N	D	A	O	A	G	K	A	I	G	S	R	L	A	V	I	T	Q	840
Oy	841	N	I	N	A	L	G	T	G	I	I	S	I	L	I	G	M	O	L	T	L	L	A	I	V	P	I	A	I	A	G	V	E	K	M	L	S	G	A	L	K	D	K	E	L	E	G	A	K	I	A	T	900		
Db	841	N	I	N	A	L	G	T	G	I	I	S	I	L	I	G	M	O	L	T	L	L	A	I	V	P	I	A	I	A	G	V	E	K	M	L	S	G	A	L	K	D	K	E	L	E	G	A	K	I	A	T	900		
Oy	901	E	A	I	E	N	R	T	V	S	L	T	R	O	K	E	Y	M	A	O	S	L	O	V	P	R	N	S	L	R	K	A	H	I	F	G	V	S	F	S	T	O	A	M	P	S	V	A	G	C	R	960			
Db	901	E	A	I	E	N	R	T	V	S	L	T	R	O	K	E	Y	M	A	O	S	L	O	V	P	R	N	S	L	R	K	A	H	I	F	G	V	S	F	S	T	O	A	M	P	S	V	A	G	C	R	960			
Oy	961	F	G	A	I	V	A	N	E	N	F	M	F	O	V	L	L	V	F	S	A	I	F	G	A	M	A	G	O	V	S	S	F	A	P	D	A	K	A	K	A	S	A	H	I	M	I	E	K	S	P	L	1020		
Db	961	F	G	A	I	V	A	N	E	N	F	M	F	O	V	L	L	V	F	S	A	I	F	G	A	M	A	G	O	V	S	S	F	A	P	D	A	K	A	K	A	S	A	H	I	M	I	E	K	S	P	L	1020		
Oy	1021	D	S	Y	S	P	H	G	L	K	P	N	L	E	G	N	T	F	N	E	V	E	Y	N	P	T	R	D	I	P	L	O	G	L	S	L	E	V	K	G	O	T	L	A	N	S	G	C	G	S	T	1080			
Db	1021	D	S	Y	S	P	H	G	L	K	P	N	L	E	G	N	T	F	N	E	V	E	Y	N	P	T	R	D	I	P	L	O	G	L	S	L	E	V	K	G	O	T	L	A	N	S	G	C	G	S	T</				

XX 05-APR-2001.
 PD 28-SEP-2000; 2000WO-US26767.
 PF 28-SEP-1999; 99US-0156510.
 PR (GENT-) GENEST CORP.
 PA Stocker PJ, Steimel-crespi DT, Crespi CL, Reif TC, Patten CJ;
 PI WPI; 2001-235373/24.
 DR N-PSDB; AAD03505.
 XX New dog P-glycoproteins (PGP) and their encoding nucleic acids, useful
 PT for determining the bioavailability of drugs and for screening for dog
 PT PGP inhibitors -
 PS Claim 17; Page 99-102; 111pp; English.
 XX The invention relates to dog P-glycoprotein (PGP) also referred
 CC as multidrug transporter (MDR1) and nucleic acids encoding them.
 CC The invention also includes fragments and biologically functional
 CC variants of dog P-glycoprotein. PGP and their nucleic acids are
 CC useful for determining the bioavailability of drugs and for
 CC screening PGP inhibitors. They are useful for the diagnosis and
 CC treatment of conditions characterised by PGP activity, by
 CC reducing or increasing PGP activity in a cell. PGP nucleic acids
 CC are used as oligonucleotide probes. Complements of PGP nucleic
 CC acids are useful as antisense oligonucleotides, to induce a PGP
 CC 'knockout' phenotype. They are used to prepare a non-human
 CC transgenic animal, which are valuable as genetic models for
 CC human diseases.
 CC The present sequence is dog P-glycoprotein (PGP) allelic variant.
 CC This sequence is also referred as Genotype B protein. The PGP
 CC enzyme functions as an efflux pump exporting small molecules
 CC across the cell membrane. This enzyme is a member of the ABC
 CC transporter family.
 CC
 XX Sequence 1281 AA:
 SX
 Query Match 99.8%; Score 6463; DB 22; Length 1281;
 Best Local Similarity 99.8%; Pred. No. 0;
 Matches 1279; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 421 SCQTVALVNGSGGCKSTTVQLMORLYDPTDGMVCIQDGIPTINVRHLREITGVVSGQPV 480
 DB 421 SGQTVALVNGSGGCKSTTVQLMORLYDPTDGMVCIQDGIPTINVRHLREITGVVSGQPV 480
 QY 481 LFATTIAENIRYGRNVTMDIEKAVKANAYDFIKLPLNKEDTLVGERGALSGGOKOR 540
 DB 481 LFATTIAENIRYGRNVTMDIEKAVKANAYDFIKLPLNKEDTLVGERGALSGGOKOR 540
 QY 541 IAIKALVRNPKILLDEATSALDTSEAVVQALDKARKGRTTIVIAHRLSTVRNADVI 600
 DB 541 IAIKALVRNPKILLDEATSALDTSEAVVQALDKARKGRTTIVIAHRLSTVRNADVI 600
 QY 601 AGPDDGVIVKGNHDELMKEKGIYFKLVMTQTRGNETELEENATGESKSSDALNEMSPDS 660
 DB 601 AGPDDGVIVKGNHDELMKEKGIYFKLVMTQTRGNETELEENATGESKSSDALNEMSPDS 660
 QY 661 GSSLIKRRSTRRSIHAPOGODRKIGTEDENENVPVPSFWRILKLKSTMPFVVGIFCA 720
 DB 661 GSSLIKRRSTRRSIHAPOGODRKIGTEDENENVPVPSFWRILKLKSTMPFVVGIFCA 720
 QY 721 IINGLOPAFSIIFRSIIIGITRDEDEPDKRONSMPVFLVIGIISFTFFLOGFTFG 780
 DB 721 IINGLOPAFSIIFRSIIIGITRDEDEPDKRONSMPVFLVIGIISFTFFLOGFTFG 780
 QY 781 KAGELITRRLRYMFRSMLRQDVSWFDDPKNTTCALITRANAAOYKKAIGSRLAVITQ 840
 DB 781 KAGELITRRLRYMFRSMLRQDVSWFDDPKNTTCALITRANAAOYKKAIGSRLAVITQ 840
 QY 841 NIANGTGIISLIGMOLTLILAIVPIIAIGAVEMKMLSGALKDKKELBGAGKIAT 900
 DB 841 NIANGTGIISLIGMOLTLILAIVPIIAIGAVEMKMLSGALKDKKELBGAGKIAT 900
 QY 901 EAIENFRVVSILTRQKEEYMAOSLOVPRNSLRKAHIGSVSSTIQAMMYSYAGCFR 960
 DB 901 EAIENFRVVSILTRQKEEYMAOSLOVPRNSLRKAHIGSVSSTIQAMMYSYAGCFR 960
 QY 961 FGAYIVANEFNMFQDVLVFSALVIGANAVQVSSFPADYAKAVSAHVIMITEKSPLI 1020
 DB 961 FGAYIVANEFNMFQDVLVFSALVIGANAVQVSSFPADYAKAVSAHVIMITEKSPLI 1020
 QY 1021 DSYSPHGLKPNLTLEGNTFENEVFNYPTRPDIPIVLQSLSEYKGGQTLALVSGSGCKST 1080
 DB 1021 DSYSPHGLKPNLTLEGNTFENEVFNYPTRPDIPIVLQSLSEYKGGQTLALVSGSGCKST 1080
 QY 1081 VVOLLERFYDPLAGSVLIDGKEIKHLNVQMLRAHLGIVSOEPILEFDCSIAENIYGDNSR 1140
 DB 1081 VVOLLERFYDPLAGSVLIDGKEIKHLNVQMLRAHLGIVSOEPILEFDCSIAENIYGDNSR 1140
 QY 1141 VVSHHEIMQAAKANIHHFLETLPKRYNTRVGDGTOLSGGOKORIAIARALVRQPHLL 1200
 DB 1141 VVSHHEIMQAAKANIHHFLETLPKRYNTRVGDGTOLSGGOKORIAIARALVRQPHLL 1200
 QY 1201 LDEATSLDTESERVOEALDKAREGRTCIYIAHRLSTIQNADLIVFONKVKVHEGHTQ 1260
 DB 1201 LDEATSLDTESERVOEALDKAREGRTCIYIAHRLSTIQNADLIVFONKVKVHEGHTQ 1260
 QY 1261 QLLAQKGIYFSMVSVOAGAKR 1281
 DB 1261 QLLAQKGIYFSMVSVOAGAKR 1281

RESULT 4
 ID AAE00310 standard; Protein: 1281 AA.
 AC AAE00310;
 AC 13-JUN-2001 (first entry)
 DE Dog P-glycoprotein (PGP) allelic variant (Genotype D).
 XX Dog; P-glycoprotein allelic variant; PGP; multidrug transporter;
 KW MDR1; drug bioavailability; transgenic animal; genetic model.
 KW

XX Canis familiaris.
OS
XX
FH Key Location/Qualifiers
FT MISC-difference 25
FT MISC-difference 197 /note= "Asn of GenotypeC substituted by Lys"
FT MISC-difference 329 /note= "His of GenotypeC substituted by Glu"
FT MISC-difference 1148 /note= "Ser of GenotypeC substituted by Thr"
FT MISC-difference 1148 /note= "Met of GenotypeC substituted by Val"
XX
PN W0200123540-A2.
XX
PD 05-APR-2001.
XX
PF 28-SEP-2000; 2000MO-US26767.
XX
PR 28-SEP-1999; 99US-0156510.
XX
PR (GENT-) GENTEST CORP.
XX
PI Stocker PJ, Stelmel-crespi DT, Crespi CL, Reif TC, Patten CJ;
DR WPI; 2001-235373/24.
XX
DR N-PSDB; AAD03506.
XX
PT New dog P-glycoproteins (PGP) and their encoding nucleic acids, useful
PT for determining the bioavailability of drugs and for screening for dog
PT PGP inhibitors -
XX
PS Claim 17; Page 108-110; 11pp; English.
XX
XX The invention relates to dog P-glycoprotein (PGP) also referred
CC as multidrug transporter (MDR1) and nucleic acids encoding them.
CC The invention also includes fragments and biologically functional
CC variants of dog P-glycoprotein. PGP and their nucleic acids are
CC useful for determining the bioavailability of drugs and for
CC screening PGP inhibitors. They are useful for the diagnosis and
CC treatment of conditions characterised by PGP activity, by
CC reducing or increasing PGP activity in a cell. PGP nucleic acids
CC are used as oligonucleotide probes. Complements of PGP nucleic
CC acids are useful as antisense oligonucleotides, to induce a PGP
CC 'knockout' phenotype. They are used to prepare a non-human
CC transgenic animal, which are valuable as genetic models for
CC human diseases.
CC The present sequence is dog P-glycoprotein (PGP) allelic variant.
CC This sequence is also referred as Genotype D protein. The PGP
CC enzyme functions as an efflux pump exporting small molecules
CC across the cell membrane. This enzyme is a member of the ABC
CC transporter family.
CC
XX
SQ Sequence 1281 AA;
Query Match 99.7%; Score 6456; DB 22; Length 1281;
Best Local Similarity 99.7%; Pred. No. 0;
Matches 1277; Conservative 2; Mismatches 2; Indels 0; Gaps 0;
QY 1 MDPBGKSGSAEKNTKMGKSKKNEKKKPYSTFAMFRYSNWLDRLYMLVGTMAII 60
DB 1 MDPBGKSGSAEKNTKMGKSKKNEKKKPYSTFAMFRYSNWLDRLYMLVGTMAII 60
QY 61 HGAALPLMLVFGNMTDSFANAGISRNKTFVIIINESTINNTQHFNIHLEEMTYAAY 120
DB 61 HGAALPLMLVFGNMTDSFANAGISRNKTFVIIINESTINNTQHFNIHLEEMTYAAY 120
QY 121 SGGAGVAAAYIOVSFNCIAGROILKIRKOFHAIHQEIGMDYVDHVGTLRLTDD 180
DB 121 SGGAGVAAAYIOVSFNCIAGROILKIRKOFHAIHQEIGMDYVDHVGTLRLTDD 180
QY 181 VSKINBGIDKIGMFPHSIAFTFGTIVGTGKMLTLVLAISPYLGSAIAWAKLISS 240
DB 181 VSKINBGIDKIGMFPHSIAFTFGTIVGTGKMLTLVLAISPYLGSAIAWAKLISS 240

DB 181 VSKINBGIDKIGMFPHSIAFTFGTIVGTGKMLTLVLAISPYLGSAIAWAKLISS 240
QY 241 FTDKELLAAVAKAGAAVEVLAIIRTVIAFGOKKLEERYNKNLEBAKIGIKKAITANIS 300
DB 241 FTDKELLAAVAKAGAAVEVLAIIRTVIAFGOKKLEERYNKNLEBAKIGIKKAITANIS 300
QY 301 IGAAPLLIYASALAFWGTSLVSSSEYIGOVLTWFESVLIGARSIGQASPSTEAFANA 360
DB 301 IGAAPLLIYASALAFWGTSLVSSSEYIGOVLTWFESVLIGARSIGQASPSTEAFANA 360
QY 361 RGAAYEIFKIIDNKSIDSYSKSHKPDNKGMLERNVHFSYPSRKVKYILKGLNKKYQ 420
DB 361 RGAAYEIFKIIDNKSIDSYSKSHKPDNKGMLERNVHFSYPSRKVKYILKGLNKKYQ 420
QY 421 SGQTVLVGNSCGSKSTVQLMQRLYDPTGDMVICIGODIRITINVRHLREITGVVSOEPV 480
DB 421 SGQTVLVGNSCGSKSTVQLMQRLYDPTGDMVICIGODIRITINVRHLREITGVVSOEPV 480
QY 481 LEFTTAENIRYGRNVTMDIEKAVKANAYDFIMKLNFKFDTLVGRGAQLSGGQKOR 540
DB 481 LEFTTAENIRYGRNVTMDIEKAVKANAYDFIMKLNFKFDTLVGRGAQLSGGQKOR 540
QY 541 IAIARALVNRPKIILDBATSALDTESBAVVOVALDKARKGRTTVIAHRLSYRANAVI 600
DB 541 IAIARALVNRPKIILDBATSALDTESBAVVOVALDKARKGRTTVIAHRLSYRANAVI 600
QY 601 AGFDDGVIEKGNHDELMKREKGIYFKLVMTQNGNETELENATGSEKSESDALEMSPRDS 660
DB 601 AGFDDGVIEKGNHDELMKREKGIYFKLVMTQNGNETELENATGSEKSESDALEMSPRDS 660
QY 661 GSSLIRKSTRSIHAPGQDRKLTGTDENENVPYPSVRILKINSEWYFVVGIFCA 720
DB 661 GSSLIRKSTRSIHAPGQDRKLTGTDENENVPYPSVRILKINSEWYFVVGIFCA 720
QY 721 IINGLOPAFSIISRIIGITFRODEPEYTRKONSNSVLFVLICISITFTPLQGTFFG 780
DB 721 IINGLOPAFSIISRIIGITFRODEPEYTRKONSNSVLFVLICISITFTPLQGTFFG 780
QY 781 KAGEIITKRLRVRNRSMLRODVSMFPDPKNTGALTTLANDAAVGAIGSLRAVITQ 840
DB 781 KAGEIITKRLRVRNRSMLRODVSMFPDPKNTGALTTLANDAAVGAIGSLRAVITQ 840
QY 841 NINMGSTGIIILYIGMQLTLLAIIVPIAINGVEMKLSGQALKKKELEAGKAT 900
DB 841 NINMGSTGIIILYIGMQLTLLAIIVPIAINGVEMKLSGQALKKKELEAGKAT 900
QY 901 EAIENFRVVSILTRQKFEYWAOSLOVPYRNSLRRAHIFGVFSITQAMMFSTYACFR 960
DB 901 EAIENFRVVSILTRQKFEYWAOSLOVPYRNSLRRAHIFGVFSITQAMMFSTYACFR 960
QY 961 FGAYLVANEFNMFODVILVFSAIVEGAMAVGOVSFSPAPYAKAKVSAHAHYIMIEKSPLI 1020
DB 961 FGAYLVANEFNMFODVILVFSAIVEGAMAVGOVSFSPAPYAKAKVSAHAHYIMIEKSPLI 1020
QY 1021 DSYSPHGLKPNLTLEGNVTENFVFNYPTRPDIPYLGGLSEYKKQOTLALVSSGCKST 1080
DB 1021 DSYSPHGLKPNLTLEGNVTENFVFNYPTRPDIPYLGGLSEYKKQOTLALVSSGCKST 1080
QY 1081 VVOLLEFRTDPLAGSVLIDGKEIKHLNVOQLRAHLCISOEPIIFDCSIAENIAYGNSR 1140
DB 1081 VVOLLEFRTDPLAGSVLIDGKEIKHLNVOQLRAHLCISOEPIIFDCSIAENIAYGNSR 1140
QY 1141 VVSHEEIMQAAKEANIHHFETLPEKYNTRVGDKGTQLSGGQKQRIAIARALVROPHILL 1200
DB 1141 VVSHEEIMQAAKEANIHHFETLPEKYNTRVGDKGTQLSGGQKQRIAIARALVROPHILL 1200
QY 1201 LDEATSLDTESEKVVQALDKAREGRTCIIVAHRLSTIONMDLIVVONKKEVHEHTQ 1260
DB 1201 LDEATSLDTESEKVVQALDKAREGRTCIIVAHRLSTIONMDLIVVONKKEVHEHTQ 1260
QY 1261 QLLAQKGIIFSMVSVQAGAKR 1281
DB 1261 QLLAQKGIIFSMVSVQAGAKR 1281

Db 180 VSKINEGIGDKVGMFFQSIATFTTGFIVGFTPGWKLLVLAILSPVLGLSAIWKILSS 239

[illegible]

```

RESULT 6
AAE00304
ID AAE00304 standard; Protein; 1280 AA.
XX
AC AAE00304;
XX
XX 13-JUN-2001 (first entry)
XX
XX Dog P-glycoprotein (PGP) #2.
XX
XX Dog: P-glycoprotein; PGP; multidrug transporter; MDR1;
XX drug bioavailability; transgenic animal; genetic model.
XX
OS Canis familiaris.
XX
XX WO200123540-A2.
XX
XX 05-APR-2001.
XX
XX 28-SEP-2000; 2000WO-US26767.
XX
XX 28-SEP-1999; 99US-0156510.
XX
XX (GENT-) GENTEST CORP.
XX
XX Stocker PJ, Steimel-crespi DT, Crespi CL, Reif TC, Patten CJ;
XX WPI; 2001-235373/24.
XX
XX N-PSDB; AAD03489.
XX
XX New dog P-glycoproteins (PGP) and their encoding nucleic acids, useful
XX for determining the bioavailability of drugs and for screening for dog
XX PGP inhibitors.
XX
XX Example 2; Page 72-75; 11pp; English.
XX
XX The invention relates to dog P-glycoprotein (PGP) also referred
XX as multidrug transporter (MDR1) and nucleic acids encoding them.
XX The invention also includes fragments and biologically functional
XX variants of dog P-glycoprotein. PGP and their nucleic acids are
XX useful for determining the bioavailability of drugs and for
XX screening PGP inhibitors. They are useful for the diagnosis and
XX treatment of conditions characterised by PGP activity, by
XX reducing or increasing PGP activity in a cell. PGP nucleic acids
XX are used as oligonucleotide probes. Complements of PGP nucleic
XX 'knockout' phenotype. They are used to prepare a non-human
XX transgenic animal, which are valuable as genetic models for
XX human diseases.
XX
XX The present sequence is dog P-glycoprotein (PGP). The
XX PGP enzyme functions as an efflux pump exporting small molecules
XX across the cell membrane. This enzyme is a member of the ABC
XX transporter family.
XX
XX Sequence 1280 AA:
XX
Query Match 99.1%; Score 6420.5; DB 22; Length 1280;
Best Local Similarity 99.3%; Pred. No. 0;
Matches 1272; Conservative 4; Mismatches 4; Indels 1; Gaps 1;
QY 1 MDEPGGKGSAAEKNFWMGKSKSKKNEKKKPTSTPFMFYRYSNNLDRLYLVTGMAAI 60
DB 1 MDEPGGKGSAAEKNFWMGKSKSKK-EKKEKPTTSTPFMFYRYSNNLDRLYLVTGMAAI 59
QY 61 HGAALPLMIVGNTMTDSFANNGISRNKTPVYIINESTITNTQHFIIHLEEMTTAYYY 120
DB 60 HGAALPLMIVGNTMTDSFANNGISRNKTPVYIINESTITNTQHFIIHLEEMTTAYYY 119
QY 121 SGIGAGVLAAYIYVSFMCIAAGROIILKIRKOFHAIIMROEIGFVDVHVGELNTRLTDD 180
DB 120 SGIGAGVLAAYIYVSFMCIAAGROIILKIRKOFHAIIMROEIGFVDVHVGELNTRLTDD 179

```

```

QY 181 VSKINEGIDKIGMFHSIAFTFGTIVGTGRMKLTIIYLAISVYLGSIAIWAIISS 240
DB 180 VSKINEGIDKIGMFHSIAFTFGTIVGTGRMKLTIIYLAISVYLGSIAIWAIISS 239
QY 241 FTDKELIAYAKAGVAEEVLAIRIVIAFGQKKELERYNKNLEBAKGIKKAITANIS 300
DB 240 FTDKELIAYAKAGVAEEVLAIRIVIAFGQKKELERYNKNLEBAKGIKKAITANIS 299
QY 301 IGAAPFLIYASVALAFWYGSILVSEYSIGOVLPVFSVYLIGAFSIGQASPIEAFANA 360
DB 300 IGAAPFLIYASVALAFWYGSILVSEYSIGOVLPVFSVYLIGAFSIGQASPIEAFANA 359
QY 361 RGAAYEIEFKIIDNKPSIDSYSGKHPDNKGNLEFFKNVHSYSPSRKREVIKGLNKKYQ 420
DB 360 RGAAYEIEFKIIDNKPSIDSYSGKHPDNKGNLEFFKNVHSYSPSRKREVIKGLNKKYQ 419
QY 421 SGOYVALVGNSSCGKSTYVQLMORLYDPTDGMVCIQODIRTIWVRLREITGVSGOEPV 480
DB 420 SGOYVALVGNSSCGKSTYVQLMORLYDPTDGMVCIQODIRTIWVRLREITGVSGOEPV 479
QY 481 LEATTIAENIRGRBNVTMDIEKAKEANAAYDEIMKLPMKPTIIVGERGAOLSGGOKOR 540
DB 480 LEATTIAENIRGRBNVTMDIEKAKEANAAYDEIMKLPMKPTIIVGERGAOLSGGOKOR 539
QY 541 IAIARALVBNPKIILLDEATSAIDTSEAVVOVALDKARKGRTTIVIAHRLSTVRNADVI 600
DB 540 IAIARALVBNPKIILLDEATSAIDTSEAVVOVALDKARKGRTTIVIAHRLSTVRNADVI 599
QY 601 AGFDGVIVVEKGNHDELMKKEGIFKLVMTQTRNGNEIELENAATESKSDALESPDS 660
DB 600 AGFDGVIVVEKGNHDELMKKEGIFKLVMTQTRNGNEIELENAATESKSDALESPDS 659
QY 661 GSSLIKRRSTRSIIHAPOGDRKLTGKEDLENVBSVSWRLIKLNSTEMPEYVVGJFCA 720
DB 660 GSSLIKRRSTRSIIHAPOGDRKLTGKEDLENVBSVSWRLIKLNSTEMPEYVVGJFCA 719
QY 721 IINGLOPASIIIFSRIGIFTRDEDEPETKRONSNMFVFLVIGIISFIFIFLOGTFG 780
DB 720 IINGLOPASIIIFSRIGIFTRDEDEPETKRONSNMFVFLVIGIISFIFIFLOGTFG 779
QY 781 KAGEIILTKRLRYVFSMLRODVSNFDDPKNTGALTTRLANDAAYKAGIAGSLAVITQ 840
DB 780 KAGEIILTKRLRYVFSMLRODVSNFDDPKNTGALTTRLANDAAYKAGIAGSLAVITQ 839
QY 841 NINMLTGIIISLITYGWOLTLILLAIVPIIAAGVVEKMLSGOALDKKELBAGKAT 900
DB 840 NINMLTGIIISLITYGWOLTLILLAIVPIIAAGVVEKMLSGOALDKKELBAGKAT 899
QY 901 EAIENFRVVSILTRQKFEYMAAGSLQVRYNSLRKAHIGVSFSITQAMMYPSVAGCFR 960
DB 900 EAIENFRVVSILTRQKFEYMAAGSLQVRYNSLRKAHIGVSFSITQAMMYPSVAGCFR 959
QY 961 FGAYIVANEFMFQOVLLVFSIAIVGAMAVGOVSSFAVDYAKAKVSAHYIMIEKSPIL 1020
DB 960 FGAYIVANEFMFQOVLLVFSIAIVGAMAVGOVSSFAVDYAKAKVSAHYIMIEKSPIL 1019
QY 1021 DSYSPHGLKPTLIGNTYFENVFNYPTRDIPYLOGLSIEVKKGOTLALVSSCGKST 1080
DB 1020 DSYSPHGLKPTLIGNTYFENVFNYPTRDIPYLOGLSIEVKKGOTLALVSSCGKST 1079
QY 1081 VOLLERYDPLAGSVLIDKEIKHLNVOMLRAHLVISOEPIILFPCSTIAENIAGVDSNR 1140
DB 1080 VOLLERYDPLAGSVLIDKEIKHLNVOMLRAHLVISOEPIILFPCSTIAENIAGVDSNR 1139
QY 1141 VVSHEEIMQAKKANIHFIETLPEKYNTRYDGKGYQLSGGOKORAIARALVROPHTL 1200
DB 1140 VVSHEEIMQAKKANIHFIETLPEKYNTRYDGKGYQLSGGOKORAIARALVROPHTL 1199
QY 1201 LDEATSAIDTESKVVQOALDKAREGRTCTIVIAHRLSTIQNADILIVQONKVEHETHQ 1260
DB 1200 LDEATSAIDTESKVVQOALDKAREGRTCTIVIAHRLSTIQNADILIVQONKVEHETHQ 1259
QY 1261 QLLAKGIIYFSMVSVOAGAKR 1281

```

DB 1260 QLLAKGITYFSMISVQAGAKR 1280

RESULT 7

AA044073 standard; protein; 1280 AA.

AA044073:

26-JUN-1998 (first entry)

Human multidrug resistance P-glycoprotein MDR1.

Human; multidrug resistance P-glycoprotein; MDR1; prokaryotic homologue.

Homo sapiens.

Key Location/Qualifiers

Region 1..640 /note="MDR1-N from Fig 1"

Region 641..1280 /note="MDR1-C from Fig 1"

W09740160-A1.

30-OCT-1997.

24-APR-1997: 97MO-N000216.

24-APR-1996: 96EP-0201094.

(UNCR-) RUKSUNIV GROMINGEN.

Bolhuis H, Konings WN, Van Veen HW, Venema K;

WPI: 1997-535844/49.

Prokaryotic homologue of human multiple drug resistance protein -
used to screen for compounds that inhibit, or avoid, drug resistance

Claim 10: Fig 1, 35pp: English.

The present invention describes a recombinant or isolated nucleic acid (1), derived from a prokaryotic gene, which encodes at least a specific and/or functional part of a transporter protein (TP), or its derivatives, which has functional and/or structural similarity with the P-glycoprotein (Pg) encoded by the human multidrug resistance (MDR)1 gene. The present sequence represents the human MDR1 protein, derived from MDR1-N and MDR1-C as shown in the specification in figure 1. (1) is used to express recombinant proteins; its fragments are also useful as probes and primers for detection and amplification of related DNA. The protein produced, or cells expressing them, are used to determine if substances can inhibit, or avoid, MDR proteins, and in a screening method for identifying compounds that inhibit transport of cytotoxic substances from cells. Also, cells with a transmembrane protein, especially where expressed from (1), can provide (additional) MDR, particularly for use as a model system to study mechanisms of action of Pg.

Sequence 1280 AA:

Query Match 90.5%; Score 5861; DB 18; Length 1280;
Best Local Similarity 90.7%; Pred. No. 0;

Matches 1163; Conservative 55; Mismatches 60; Indels 4; Gaps 4;

DB 1 MDPGSGRGS-A-EKNFMKSKSKRNEKKKPYSTFAMFRYSNNLDRPLMYLGMMAI 59
1 MDLSDRNGGKKNKPKLNKSKER-DKREKPYVSFSMRYSNMLDKLIMVAGLAAI 59
QY 60 IHGAGLPLMLLVFGNMTDSFANAGISRNKTEPVIIINESTNTNQHFINLSEMTYAVY 119
DB 60 IHGAGLPLMLLVFGNMTDSFANAGISRNKTEPVIIINESTNTNQHFINLSEMTYAVY 117

QY 120 YSGIGAGLVAAAYIOVSFWCLAGROILKIRKOFFHAIMROEIGMDVHDVDELNTRLTD 179
DB 118 YSGIGAGLVAAAYIOVSFWCLAGROILKIRKOFFHAIMROEIGMDVHDVDELNTRLTD 177
QY 180 DYSKINGIGDKIGMFPHSIAFTFTGFTVGFTRGMKLTLYILAIISPVLAGSAIYNAKLIS 239
DB 178 DYSKINGIGDKIGMFPHSIAFTFTGFTVGFTRGMKLTLYILAIISPVLAGSAIYNAKLIS 237
QY 240 SFTDKELLATAKAGAAVEEYLAIRTVIAFGGOKKELEERNNKULEBAKIGIKKAITANI 299
DB 238 SFTDKELLATAKAGAAVEEYLAIRTVIAFGGOKKELEERNNKULEBAKIGIKKAITANI 297
QY 300 SIGAFLIYASVALAFMYGTSLVLSSEYSIGQVLTVFFSVLGAFSIGQASPSIEAFAN 359
DB 298 SIGAFLIYASVALAFMYGTSLVLSSEYSIGQVLTVFFSVLGAFSIGQASPSIEAFAN 357
QY 360 ARGAAVEIFPKIIDNKPISIDYSKSGHKPDNIKGNLEFRNNVFSYPSRKEVYIKGLNKKY 419
DB 358 ARGAAVEIFPKIIDNKPISIDYSKSGHKPDNIKGNLEFRNNVFSYPSRKEVYIKGLNKKY 417
QY 420 QSGOTVALVGNSSCGKSTTVOLMORLJPTDGMVCIDGODIRITNNRHLREITGVYSDP 479
DB 418 QSGOTVALVGNSSCGKSTTVOLMORLJPTDGMVSVGDIDIRINRFLREITGVYSDP 477
QY 480 VLFATTTAENIRYGRBNVTMDIEKAVKEANAAYDFIMKLPNKEDTLVGBRGAQLSGGOKQ 539
DB 478 VLFATTTAENIRYGRBNVTMDIEKAVKEANAAYDFIMKLPNKEDTLVGBRGAQLSGGOKQ 537
QY 540 RIATARALVBNPFIILDDPATSDTESEAVVOVALDKARKGRTTVIAHRISTYRNADV 599
DB 538 RIATARALVBNPFIILDDPATSDTESEAVVOVALDKARKGRTTVIAHRISTYRNADV 597
QY 600 IAGFDGVIYEKNDHLEKKEGIFKLVYQWOTGNIELEANAATGSKSESALMSKPD 659
DB 598 IAGFDGVIYEKNDHLEKKEGIFKLVYQWOTGNIELEANAATGSKSESALMSKPD 657
QY 660 GSSSLIKRSTRSISIAPOGODRKLGTREDLENENVPVPSFWRILKLNSTEMPYFVVGJFC 719
DB 658 SRSLSIKRSTRSISIAPOGODRKLGTREDLENENVPVPSFWRILKLNSTEMPYFVVGJFC 717
QY 720 AINGGLOPAFSLIFSRIGITRDEPETKRONSNFVLFIQISFTFFLOGTF 779
DB 718 AINGGLOPAFSLIFSRIGITRDEPETKRONSNFVLFIQISFTFFLOGTF 777
QY 760 GKAGEIILTKRLRYMVRSMLRDVSFDDPKNTGALTTRFLANDAAQVGAIGSLAVYT 839
DB 778 GKAGEIILTKRLRYMVRSMLRDVSFDDPKNTGALTTRFLANDAAQVGAIGSLAVYT 837
QY 840 ONIANLGTGIIISLIYGMOLTLTLLAIVPIIAIGVEMKLSGQALKDKKLEBAGKRIA 899
DB 838 ONIANLGTGIIISLIYGMOLTLTLLAIVPIIAIGVEMKLSGQALKDKKLEBAGKRIA 897
QY 900 TEALENRTVVSILTRQKREYVYAOISLOVPRNSLRKAHIFGFSFITOMMYFYACCF 959
DB 898 TEALENRTVVSILTRQKREYVYAOISLOVPRNSLRKAHIFGFSFITOMMYFYACCF 957
QY 960 RFGAIVLANEFMFOVLLVESAIYFGMAVGOVSSAPDIYAAKASAAHVIMITEKPL 1019
DB 958 RFGAIVLANEFMFOVLLVESAIYFGMAVGOVSSAPDIYAAKASAAHVIMITEKPL 1017
QY 1020 IDSYPHGLKPNLLEGNVTFNEVFNPTRPDIPLVLOGLSLEVKKGQTLALVSSGCGKS 1079
DB 1018 IDSYPHGLKPNLLEGNVTFNEVFNPTRPDIPLVLOGLSLEVKKGQTLALVSSGCGKS 1077
QY 1080 TVYOLLERFYPDLASVCLLDGKEIKHNLVOMLRHAGIYQOEPIIPDGSIAENIAYGNS 1139
DB 1078 TVYOLLERFYPDLASVCLLDGKEIKHNLVOMLRHAGIYQOEPIIPDGSIAENIAYGNS 1137
QY 1140 RVYSHEEIMQAKAENIHFIEFTPEKYNTRVGDGKTQLSGGOKRIATARALVROPHIL 1199
DB 1138 RVYSHEEIMQAKAENIHFIEFTPEKYNTRVGDGKTQLSGGOKRIATARALVROPHIL 1197

Qy	1200	LDENATSDLTDESEKVVQALDRAKRGCTCIVIAHRLSTIQNDLIVFQNGKVEHGTH	1259
Db	1198	LIDENATSDLTDESEKVVQALDRAKRGCTCIVIAHRLSTIQNDLIVFQNGKVEHGTH	1259
Qy	1260	QQLIAQGIYFSVWVQAGAKR	1281
Db	1258	QQLIAQGIYFSVWVQAGAKR	1279
RESULT 8			
AAV58186			
ID	AAV58186	standard; Protein: 1280 AA.	
AC			
XX	AAV58186;		
XX			
DT	14-MAR-2000	(first entry)	
DX			
DE	Human wild-type multidrug resistance-1 (MDR-1) protein.		
XX			
KM	Multidrug resistance: MDR-1; P-glycoprotein;		
KM	transmembrane efflux pump; hematopoietic stem cell; transduction;		
KM	bone marrow transplantation; chemotherapy; radiation therapy; cancer;		
KW	gene therapy; gene replacement; genetic defect; thalassemia;		
KW	Gaucher's disease; sickle cell anaemia; leukaemia; ex vivo expansion;		
KW	cytokine; wild-type.		
XX			
OS	Homo sapiens.		
XX			
FH	Key	Location/Qualifiers	
FT	Misc-difference 185	/note="This residue is Val in a mutant MDR-1	
FT		(AAV58187)"	
XX			
PN	W09961589-A2.		
XX			
PD	02-DEC-1999.		
XX			
PF	27-MAY-1999;	99WO-US11825.	
XX			
PR	28-MAY-1998;	98US-0086988.	
XX			
PA	(SJUD-) ST JUDE CHILDREN'S RES HOSPITAL.		
XX			
PI	Sorrentino B, Bunting K;		
XX			
DR	WPI: 2000-072615/06.		
XX	N-PSDB; AAZ49332.		
XX			
PT	Ex vivo expansion of hematopoietic stem cells transduced with a		
PT	sequence encoding human multidrug resistance-1, used for bone marrow		
PT	transplantation -		
XX			
PS	Claim 10; Page 71-79; 113pp; English.		
XX			
CC	This sequence represents human wild-type multidrug		
CC	resistance protein MDR-1. MDR-1 is a transmembrane efflux		
CC	pump, responsible for the export of drugs from certain		
CC	cells, particularly cancer cells. Wild-type MDR-1 shows		
CC	increased resistance to etoposide and decreased resistance to vinca		
CC	alkaloids compared with a mutant form (AAV58187) where the Gly at		
CC	position 185 is replaced by Val. The invention relates to transducing		
CC	hematopoietic stem cells with nucleic acid encoding an MDR protein		
CC	and culturing the modified cells. The modified hematopoietic stem		
CC	cells are useful in bone marrow transplantation (to reconstitute		
CC	hematopoietic systems in patients who have undergone chemotherapy or		
CC	radiation therapy) and in ex vivo gene therapy of genetic defects in		
CC	cells derived from hematopoietic stem cells, e.g., thalassemia,		
CC	Gaucher's disease, sickle cell anaemia or leukaemia. The modified		
CC	cells can also be used to identify factors involved in regulating		
CC	proliferation and differentiation in hematopoietic stem cells.		
CC	Hematopoietic stem cells that express MDR-1 will be protected against		
CC	chemotherapeutic agents, so can be engrafted while the patient is		
CC	undergoing chemotherapy. Expansion of (rare) hematopoietic stem cells		

CC	provides sufficient cells to permit standard biochemical analysis.
CC	Overexpression of MDR-1 allows cytokine-driven expansion of
CC	hematopoietic stem cells by at least 10-fold compared with a maximum
CC	of 4-fold in known procedures.
XX	
50	Sequence 1280 AA:
Query Match	90.5%; Score 5861; DB 21; Length 1280;
Best Local Similarity	90.7%; Pred. No. 0;
Matches 1163; Conservative 55; Mismatches 60; Indels 4; Gaps 4;	
QY	1 MDEPGGKRGSA-EKNEFKMGKRSKKNEKKKKRTVTSEFAMFRSNMLDLNLYGTMAI 59
Db	1 MDLEGDNGAKAKKNPFKLNNKSEK-DKEKKPTYSVFSEMRISNLDLYNVTGLAAI 59
QY	60 IHGALPLMLLVFGNMTDSFANAGISRNTFPYINESTTNTTQHFNLLEEMETVAY 119
Db	60 IHGAGLPMLLVGEEMTDIFANAG-NLEDMNSITNRSDINDTGFPM-LEEDMIRVAY 117
QY	120 YSGIGAGVLVAAYIYVSPFCIAGROILIKRQFPHAIKRQIGFVDVHDELTRLTD 179
Db	118 YSGIGAGVLVAAYIYVSPFCIAGROILIKRQFPHAIKRQIGFVDVHDELTRLTD 177
QY	180 DYSKINIEGIDKIGMFEHSIATFETFTFYGIFGRMKTLVLIAISPVYGLSAIAKILTS 239
Db	178 DYSKINIEGIDKIGMFEHSMATFETFTFYGIFGRMKTLVLIAISPVGLSAAVAKILTS 237
QY	240 SFPDKRLAYAKAGNAEVLAIKRVIAFGQKKELERYKNLDEAKGIGIKAITANI 299
Db	238 SFPDKRLAYAKAGNAEVLAIKRVIAFGQKKELERYKNLDEAKRIGIKAITANI 297
QY	300 SIGAPELLIYASALAFWGTSLVLSSEYSIGOVLTFFSVLIGAFSIGQASPTIEAFAN 359
Db	298 SIGAPELLIYASALAFWGTTLVLSGEYSIGOVLTFFSVLIGAFSIGQASPTIEAFAN 357
QY	360 AGGAAEIKITIDNKPISDYSKSHKPDNIKGLNEFKNVHSPSRKREVKILGLNLKV 419
Db	358 AGGAAEIKITIDNKPISDYSKSHKPDNIKGLNEFKNVHSPSRKREVKILGLNLKV 417
QY	420 OSGQVYALVGNSSGCKSTVOLMQRLPYPTQMGVINDQDRIITNVRILREITVVSQEP 479
Db	418 OSGQVYALVGNSSGCKSTVOLMQRLPYPTQMGVINDQDRIITNVRILREITVVSQEP 477
QY	480 VLFATTIAENIRYGENVTMDIEIKAVKANAYDFIMKLPNKFPTLVGERGAQISGGQK 539
Db	478 VLFATTIAENIRYGENVTMDIEIKAVKANAYDFIMKLPNKFPTLVGERGAQISGGQK 537
QY	540 RLAIKRALVVRNPKLILLDLEATSALDTESEAAVQVALDKARKGRITTYIAHRLSTVRNAV 599
Db	538 RLAIKRALVVRNPKLILLDLEATSALDTESEAAVQVALDKARKGRITTYIAHRLSTVRNAV 597
QY	600 IAGFDGVIVKGNHDELMKEKGIYFKLYVTQTRGNEIELENATGESKESDLEMPKD 659
Db	598 IAGFDGVIVKGNHDELMKEKGIYFKLYVTQTRGNEIELENADSEKSIDLEMPKND 657
QY	660 SSSSLIKRRSRRRSIHAQGGODRKLGRKEDLNEVPPVSWRILILKNSTWPPYFVVGIFC 719
Db	658 SSSSLIKRRSRRRSRQAOQDRKLSTKEALDESIPVSWRIMKILNLTWPPYFVVGIVC 717
QY	720 AIIINGGLQPAISIIFSRIIGITRDEDEPERKROSNNEVSFLVYLIGISITFELGGTFE 779
Db	718 AIIINGGLQPAVAIIFSKIIIGITRIDDEPERKROSNNLSILLFLALGISITTFELGGTFE 777
QY	780 GKAEIILKRLIRYVFRFSMLRQDYSWPDDPKNTTGALTRLANDAAQVGAIGSRILAVIT 839
Db	778 GKAEIILKRLIRYVFRFSMLRQDYSWPDDPKNTTGALTRLANDAAQVGAIGSRILAVIT 837
QY	840 QNINLGTGIIISLIYQMQLTLLILAVPIITAIAGVEMKMSGOALKRKELEGAGKTA 899
Db	838 QNINLGTGIIISLIYQMQLTLLILAVPIITAIAGVEMKMSGOALKRKELEGAGKTA 897
QY	900 TEALENRTYVSLTRCKFEYMTAQSLOVPRNSLRKAHIFGVSESTITQAMMTFSYAGCF 959

Db 898 TEAIEHFVTVSLTREQFEHMYAQSLOVPRNSLRKAHIGITFSFOAMMYFVAGCF 957
QY 960 REGAYLVANEFNFOVDLVLSAIVGAMAVGOVSPAPDYAKKVAHAHVIIMIEKSP 1019
Db 958 REGAYLVANEFNFOVDLVLSAIVGAMAVGOVSPAPDYAKKVAHAHVIIMIEKSP 1017
QY 1020 IDSYPHGLKPNTEGNTFENEVFNYPTRPDIPVLOGLSIEVKKGOTLALVSSGGCKS 1079
Db 1018 IDSYPHGLKPNTEGNTFENEVFNYPTRPDIPVLOGLSIEVKKGOTLALVSSGGCKS 1077
QY 1080 TVVOLLEREPYDLAAGVILIDGKEIKHLNVOMIRAHLLGIVSOEPILFDCSIENIAYGDN 1139
Db 1078 TVVOLLEREPYDLAAGVILIDGKEIKHLNVOMIRAHLLGIVSOEPILFDCSIENIAYGDN 1137
QY 1140 RYVSHHEIQAQKANIHHITETLPEKYNTRYGDKGTOLSGGOKRIARALVROPHIL 1199
Db 1138 RYVSHHEIQAQKANIHHITETLPEKYNTRYGDKGTOLSGGOKRIARALVROPHIL 1197
QY 1200 LIDETSALDTESEKVVQBALDKAREGRTCIYIAHRLSTIONADLIIVFONGKYKEGHT 1259
Db 1198 LIDETSALDTESEKVVQBALDKAREGRTCIYIAHRLSTIONADLIIVFONGKYKEGHT 1257
QY 1260 QOLLAOKGIYFSMVSVQAGAKR 1281
Db 1258 QOLLAOKGIYFSMVSVQAGAKR 1279

RESULT 9
AAB81959
ID AAB81959 standard; protein; 1280 AA.
XX AAB81959;
AC
DT 02-JUL-2001 (first entry)
XX
DE Human MDRL.
XX
KW Human; MDRL; multi-drug resistance; cholesterol uptake;
KW hypercholesterolemia; hypocholesterolemia; atherosclerosis;
KW coronary artery disease; cerebral vascular disease.
XX
OS Homo sapiens.
XX
PN W0200121762-A2.
XX
PD 29-MAR-2001.
XX
PF 22-SEP-2000; 2000MO-US26099.
XX
PR 23-SEP-1999; 99US-0155819.
XX
PA (BARN-) BARNES-JEWISH HOSPITAL.
XX
PI Stenson WF, Tessner T;
XX
PI WPI: 2001-328100/34.
XX
DR Modulating cellular cholesterol uptake for treating hyper
XX
PT hypocholesterolemia, involves administering an agent that inhibits or
XX
PT increases the expression of multiple drug resistance-1
XX
PS Example 1: Page 43-47; 47pp; English.
XX
CC The present invention describes a method of modulating cholesterol uptake
CC in cells, involving administering an agent capable of inhibiting the
CC multi-drug resistance protein MDRL. This is useful in the prevention and
CC treatment of cholesterol-related diseases, including
CC hypercholesterolemia, hypocholesterolemia, atherosclerosis, coronary
CC artery disease and cerebral vascular diseases. The present sequence is
CC the human MDRL protein.
XX
SQ Sequence 1280 AA;

Query Match 90.58; Score 5861; DB 22; Length 1280;
Best Local Similarity 90.78; Pred. No. 0;
Matches 1163; Conservative 55; Mismatches 60; Indels 4; Gaps 4;
QY 1 MDPEGRKGS-A-EKNFMKGGKSKKNEKKEKPPVSTFAMFRYSNMIDRLVLMVGTMAI 59
Db 1 MDPEGRKGS-A-EKNFMKGGKSKKNEKKEKPPVSTFAMFRYSNMIDRLVLMVGTMAI 59
QY 60 IHGAALPLMLVFGMTDSFANAGISRNKTPPVYINSSITNNQOHFNHLEEMTYAAY 119
Db 60 IHGAALPLMLVFGMTDSFANAGISRNKTPPVYINSSITNNQOHFNHLEEMTYAAY 117
QY 120 YSGIGAGVVAAYIOVSWRCALAGROILKIRKOFPHAIMROEIGMPDVHDELNRLTD 179
Db 118 YSGIGAGVVAAYIOVSWRCALAGROILKIRKOFPHAIMROEIGMPDVHDELNRLTD 177
QY 180 DYSKINEGIGDKIGMFHSIATFETGTVGTGKMLTVLTAISPVLGSAAMATILS 239
Db 178 DYSKINEGIGDKIGMFHSIATFETGTVGTGKMLTVLTAISPVLGSAAMATILS 237
QY 240 SFTDKELALYAKAGAAVEVLAARTVTAFGOKKELERKNLEAKGIGIKKAITANI 299
Db 238 SFTDKELALYAKAGAAVEVLAARTVTAFGOKKELERKNLEAKGIGIKKAITANI 297
QY 300 SIGAFLILYASALAFWYGTSLVLSSEYSIGOVLTVEFVSLIGAFSIGQSPSIEAFAN 359
Db 298 SIGAFLILYASALAFWYGTSLVLSSEYSIGOVLTVEFVSLIGAFSIGQSPSIEAFAN 357
QY 360 ARGAAVEIFKIIDNKPISIDYSKSGHKPDNIKGNLEFNKVFSPSRREVILKGLMKV 419
Db 358 ARGAAVEIFKIIDNKPISIDYSKSGHKPDNIKGNLEFNKVFSPSRREVILKGLMKV 417
QY 420 QSGQTVALVNSCGCKSTTVQLMQRLYDPTDGMCIDODIRITINVRHLREITGVSOEP 479
Db 418 QSGQTVALVNSCGCKSTTVQLMQRLYDPTDGMCIDODIRITINVRHLREITGVSOEP 477
QY 480 VLPATTIAENIRYGRNVMTDEIEKAVKANAAYDTIMKIPHEFDVLGERGQSLGGKQ 539
Db 478 VLPATTIAENIRYGRNVMTDEIEKAVKANAAYDTIMKIPHEFDVLGERGQSLGGKQ 537
QY 540 RIATARLVNRPKILLDETSALDTESEAVVQVALDKARKGRTIYIAHRLSTVRNADV 599
Db 538 RIATARLVNRPKILLDETSALDTESEAVVQVALDKARKGRTIYIAHRLSTVRNADV 597
QY 600 IAGFDGVIEKGNHDELMKEKGIYFKLVYTWOTRGNEIELENAATGESSESALEMSPO 659
Db 598 IAGFDGVIEKGNHDELMKEKGIYFKLVYTWOTRGNEIELENAATGESSESALEMSPO 657
QY 660 SGSSLIRKSTRSIAHPOQODRKLTGKEDLNENVPVPSFMRILKUNSTEMPYFVAGIFC 719
Db 658 SGSSLIRKSTRSIAHPOQODRKLTGKEDLNENVPVPSFMRILKUNSTEMPYFVAGIFC 717
QY 720 AIIINGLOPAFISIFRSIIIGTRDEDPETKRONSMPFVLVLVIGIISFITFLOGTF 779
Db 718 AIIINGLOPAFISIFRSIIIGTRDEDPETKRONSMPFVLVLVIGIISFITFLOGTF 777
QY 780 GKAGEILTKRLRYWFRSMRLROVSWFDDPKNTGALTTRRLANDAAQVGAIGSLAVIT 839
Db 778 GKAGEILTKRLRYWFRSMRLROVSWFDDPKNTGALTTRRLANDAAQVGAIGSLAVIT 837
QY 840 ONINNLGIGIISIIYGMOLTLTLLAVPTIATAGVEMKMSLGGALDKKKELEGAGKIA 899
Db 838 ONINNLGIGIISIIYGMOLTLTLLAVPTIATAGVEMKMSLGGALDKKKELEGAGKIA 897
QY 900 TEAIEHFVTVSLTREQFEHMYAQSLOVPRNSLRKAHIGVSEITQAMMYFVAGCF 959
Db 898 TEAIEHFVTVSLTREQFEHMYAQSLOVPRNSLRKAHIGVSEITQAMMYFVAGCF 957
QY 960 REGAYLVANEFNFOVDLVLSAIVGAMAVGOVSPAPDYAKKVAHAHVIIMIEKSP 1019
Db 958 REGAYLVANEFNFOVDLVLSAIVGAMAVGOVSPAPDYAKKVAHAHVIIMIEKSP 1017
QY 1020 IDSYPHGLKPNTEGNTFENEVFNYPTRPDIPVLOGLSIEVKKGOTLALVSSGGCKS 1079

|||||
Db 1018 IDSTSTEGSLMPNTLEGENVFENYPTRPDIPLVQLSLSTSEVKKGQFLALVSSGCGKS 1077
QY 1080 TVVOLLIERFYDPLAGSVLIDGKEIKHLNWMOMRAHLGIVSEPIILFPCOSTAENIAYGDNS 1139
Db 1078 TVVOLLIERFYDPLAGSVLIDGKEIKRLNWMOMRAHLGIVSEPIILFPCOSTAENIAYGDNS 1137
QY 1140 RVVSHBEIMQAQKANEINHHFTETLPEKYNTRVGDGTQLSGGQKORIAIARALYRPHIL 1199
Db 1138 RVVSOEIVRAKANEINHAFTESLPNKYSTKVGDKGTQLSGGQKORIAIARALYRPHIL 1197
QY 1200 LIDETSALDPTSEKVVYVADLKDAREGRTGIVIAHRLSTIQNADLIYVFGNGVKEHGT 1259
Db 1198 LIDETSALDPTSEKVVYVADLKDAREGRTGIVIAHRLSTIQNADLIYVFGNGVKEHGT 1257
QY 1260 QOLLAQKGYFSMVSVOGAKR 1281
Db 1258 QOLLAQKGYFSMVSVOGAKR 1279
RESULT 10
AAE18964
ID AAE18964 standard; Protein; 1280 AA.
AC AAE18964;
XX
DT 21-MAY-2002 (first entry)
XX
DE Human wild type P-glycoprotein (wt P-gp).
XX
KW Human; P-glycoprotein; vaccine; tumour; chemotherapeutic agent; cancer;
XX haematopoietic cell; cytotoxicity; cytosolic; P-gp.
OS Homo sapiens.
XX MO200210205-A2.
PN 07-FEB-2002.
XX
PF 01-AUG-2001; 2001MO-US24560.
XX
PR 01-AUG-2000; 2000US-222313P.
XX
PA (UNIL) UNIV ILLINOIS FOUND.
PI Ruth A, Roninson I;
XX
DR WPI; 2002-206182/26.
XX
PT New human P-glycoprotein mutants for treating cancer, comprises an
PT ability to confer increased resistance to chemotherapeutic drugs
PT relative to wild type P-glycoprotein or P-glycoprotein having a glycine
PT to valine substitution
XX
PS Claim 44; Fig 3; 47pp; English.
XX
XX The present invention relates to an isolated mutant human P-glycoprotein
XX molecule. The invention is used as vaccine. An antibody specific for the
XX mutant P-glycoprotein is useful for determining whether a human tumour
XX sample comprises tumour cells expressing mutant human P-glycoprotein. A
XX recombinant expression construct comprising nucleic acid encoding mutant
XX P-glycoprotein is useful for administering an increased amount of a
XX chemotherapeutic agent to an individual with cancer, by transducing ex
XX vivo haematopoietic cells from the individual with cancer, with the
XX expression construct, reintroducing the transduced haematopoietic cells
XX into the individual and administering an increased amount of a
XX chemotherapeutic agent without consequent haematopoietic cytotoxicity.
XX The mutant P-glycoprotein is useful for treating and diagnosing cancer
XX and for screening compounds for the capacity to disrupt binding of
XX cytotoxic drugs with the mutant P-glycoprotein or to disrupt drug efflux
XX from cells expressing the mutant P-glycoprotein. The present sequence is
XX human wild type P-glycoprotein (wt P-gp).

Seq Sequence 1280 AA:
Query Match 90.5%; Score 5861; DB 23; Length 1280;
Best Local Similarity 90.7%; Pred. No. 0;
Matches 1163; Conservative 55; Mismatches 60; Indels 4; Gaps 4;
QY 1 MDPGGRGSA-EKNFMKMGKSKKNEKKEKPPVSTPFAFRYSNMIDRLYMLVGTMAAT 59
Db 1 MLEEDRNGGAKKKNFFLNKSEK-DKEKKPPVSVSFMSFRYNMIDKLYMVGVTAAI 59
QY 60 IHGAALPLMMIVFGNMTDSFANAGISRKKTPPVVINSINNNHOHFNHLEETTYAY 119
Db 60 IHGAALPLMMIVFGNMTDIPANAG-NLEDMSNTNNDINDGFEFN-LEEDMTYAY 117
QY 120 YSGIAGVLAAYIYQVSWFCLAGROILKTRKOPFHAIMROEIGMPDVHDELNRLD 179
Db 118 YSGIAGVLAAYIYQVSWFCLAGROILKTRKOPFHAIMROEIGMPDVHDELNRLD 177
QY 180 DYSKINEGIGDKIGMFFSTATFEFTGFTVGTTRKMKLTLYIALISPYLGSAAIMAKIIS 239
Db 178 DYSKINEGIGDKIGMFFQSMATFEFTGFTVGTTRKMKLTLYIALISPYLGSAAIMAKIIS 237
QY 240 SEFDKELLAYAKAGAAVEVLAIRTVIAGGOKKELERYNKNLEKAGIGIKKAITANI 299
Db 238 SEFDKELLAYAKAGAAVEVLAIRTVIAGGOKKELERYNKNLEKAGIGIKKAITANI 297
QY 300 SIGAFLIITYASTALATWYGTSLVLSSEYSIGOVLYVFSVLIGAFSIGQASPSIBAFAN 359
Db 298 SIGAFLIITYASTALATWYGTSLVLSSEYSIGOVLYVFSVLIGAFSIGQASPSIBAFAN 357
QY 360 ARGAAVEIFKIIDKPSIDYSKSGHKPDNIKGNLEKKNHFEYPSRKVKYILKGLNLY 419
Db 358 ARGAAVEIFKIIDKPSIDYSKSGHKPDNIKGNLEKKNHFEYPSRKVKYILKGLNLY 417
QY 420 QSGQTVALVNSGCGKSTVQLMQRLYDPTDGNVCIDGODIRTNVRLHREITGVSOEP 479
Db 418 QSGQTVALVNSGCGKSTVQLMQRLYDPTDGNVCIDGODIRTNVRLHREITGVSOEP 477
QY 480 VLFATTTAENIRYGRNVYMDIEKAVKENAYDITMKLPKFFDLVGERGALSGGOK 539
Db 478 VLFATTTAENIRYGRNVYMDIEKAVKENAYDITMKLPKFFDLVGERGALSGGOK 537
QY 540 RIATARALVNRPKILLDEATSLDTESEAVVOVALDKRKRRTTVIAHRLSTVRNDV 599
Db 538 RIATARALVNRPKILLDEATSLDTESEAVVOVALDKRKRRTTVIAHRLSTVRNDV 597
QY 600 IAGFDGVIVEKGNHDELMEKGIYFKLVYMTQNGNIELENATGESKSESDALEMSPKD 659
Db 598 IAGFDGVIVEKGNHDELMEKGIYFKLVYMTQNGNIELENATGESKSESDALEMSPKD 657
QY 660 SGSSLIKRSTRSIAAPQGDOKLGTREDLNNVPVPSFRIKLNSTEMPYVVGIFC 719
Db 658 SGSSLIKRSTRSIAAPQGDOKLGTREDLNNVPVPSFRIKLNSTEMPYVVGIFC 717
QY 720 AIIINGLOPAFSIIFSRIGITFTRDEDPETKRONSMFSVLVYGIISFTFFLOGTTF 779
Db 718 AIIINGLOPAFSIIFSKIGVETRIDDEPTKRONSNLSLFLALGIIISFTFFLOGTTF 777
QY 780 GKAGELLTRLRKMYRSMRLQDVSWFDDPKTGTALTRLANDAAOVYKAGISRLAVT 839
Db 778 GKAGELLTRLRKMYRSMRLQDVSWFDDPKTGTALTRLANDAAOVYKAGISRLAVT 837
QY 840 ONIANLGTGIIISLYGMQUTLLLAIVPIIAGVVMKMLSGQALDKKELGACKIA 899
Db 838 ONIANLGTGIIISFYGMQUTLLLAIVPIIAGVVMKMLSGQALDKKELGACKIA 897
QY 900 TEAIENFRTVSLTBQKFEYMYAOSLOVPYRNSLRKAHIFVSFSTQAMMYSYAGCF 959
Db 898 TEAIENFRTVSLTBQKFEYMYAOSLOVPYRNSLRKAHIFVSFSTQAMMYSYAGCF 957
QY 960 RFGATVYANEFNPDVLLVPSAIYFGMAVGOVSSFPADYAKKVAHAHYIMITEKSP 1019
Db 958 RFGATVYANEFNPDVLLVPSAIYFGMAVGOVSSFPADYAKKVAHAHYIMITEKSP 1017

QY 1020 IDSYPHGLKPTLESGNTFENYVFNPTRPDIPVLOGLSLEVKGGOTLALVSSGGCKS 1079
DB 1018 IDSSTEGMLPTLGNNTFGVFNPTRPDIPVLOGLSLEVKGGOTLALVSSGGCKS 1077
QY 1080 IYVOLLERFYDPLAGSVLDGKEIKHLNWMLRAHLGIVSOEPIILFDCSIAENIAYGDN 1139
DB 1078 IYVOLLERFYDPLAGSVLDGKEIKHLNWMLRAHLGIVSOEPIILFDCSIAENIAYGDN 1137
QY 1140 RYVSHETIQAANKANINHHFTTLEPEKNTRVGDKGQTLSSGOKORIALANALROPHIL 1199
DB 1138 RYVSOEIEIYRAKANKANINHHFTTLEPEKNTRVGDKGQTLSSGOKORIALANALROPHIL 1197
QY 1200 LIDEATSALDTESEKVOEALDKAREGRTCIYIAHRLSTIONADLIYVFGNGKVEGHTH 1259
DB 1198 LIDEATSALDTESEKVOEALDKAREGRTCIYIAHRLSTIONADLIYVFGNGKVEGHTH 1257
QY 1260 QOLLAQGIYFSMVSVQAGAKR 1281
DB 1258 QOLLAQGIYFSMVSVQAGAKR 1279

RESULT 11
ABB07266
ID ABB07266 standard; Protein; 1280 AA.
XX
AC ABB07266;
XX
DT 26-MAR-2002 (first entry)
XX
DE Human BCRP protein related seq Id No. 2.
XX
KM Stem cell; ATP transport protein; ATP-binding cassette; antiparkinsonian;
KM hepatotopic; neurodegenerative; cytosolic; antianemic; muscular; BCRP;
KM cardiant; gene therapy.
XX
OS Homo sapiens.
XX
PN WO200192877-A2.
XX
PD 06-DEC-2001.
XX
PF 30-MAY-2001; 2001MO-US17459.
XX
PR 31-MAY-2000; 2000US-0584586.
PR 29-MAY-2001; 2001US-0866866.
XX
PA (STUD-) ST JUDE CHILDREN'S RES HOSPITAL.
XX
PI Sorrentino B, Schuetz J;
XX
DR WPI: 2002-114368/15.
DR N-PSDB: ABA94365.
XX
PT Identifying a stem cell, for treating e.g., muscular dystrophy,
PT myocardial infarction, parkinson's disease, or neurodegenerative
PT disorders, comprises detecting the expression of an ATP transport
PT protein (BCRP) by a cell
XX
PS Disclosure; Page 55-58; 87pp; English.
XX
XX The invention provides a method of identifying and/or isolating a stem
XX cell that involves detecting the expression of an ATP transport protein
XX containing a conserved ATP-binding cassette (BCRP) by a cell in a sample
XX comprising stem cells. The isolated stem cells may be used in the
XX treatment of diseases such as muscular dystrophy, degenerative liver
XX disorder, myocardial infarction, parkinson's disease, degenerative
XX disorders of the brain, and for tissue regeneration or replacement.
XX Haematopoietic cells can be used in bone marrow transplants (e.g., for
XX treatment of leukemia) and for ex vivo gene therapy for treating blood
XX diseases such as sickle cell anemia and thalassemia. The stem cells can
XX also be used as cell targets in gene therapy protocols. The present
XX sequence represents a sequence related to the BCRP for which no relevant

CC Information has been provided in the specification.
XX
SQ Sequence 1280 AA;
Query Match 90.5%; Score 5861; DB 23; Length 1280;
Best Local Similarity 90.7%; Pred. No. 0;
Matches 1163; Conservative 55; Mismatches 60; Indels 4; Gaps 4;
QY 1 MDPEGGRGSA-EKNEFKMGKSKSKNEKKEKPPVSPFAMFRYSNNDRLYMVGMAAI 59
DB 1 MDPEGGRGSA-EKNEFKMGKSKSKNEKKEKPPVSPFAMFRYSNNDRLYMVGMAAI 59
QY 60 IHCALPLMLVFGMTQSPFANAGISRNKTPPVIIINSITNTOHFTNLHEEETVYAY 119
DB 60 IHCALPLMLVFGMTQSPFANAGISRNKTPPVIIINSITNTOHFTNLHEEETVYAY 117
QY 120 YSGIGAGVLVAAYIQVSWFVLAAGHQLIKRQFFHAIIMROETWEPVHYGELNTRLT 179
DB 118 YSGIGAGVLVAAYIQVSWFVLAAGHQLIKRQFFHAIIMROETWEPVHYGELNTRLT 177
QY 180 DYSKINIEGIGDKIGMFPHSIAFEFTGFTVGRGKTLVLAISPVIGLSAAIMAKILS 239
DB 178 DYSKINIEGIGDKIGMFPHSIAFEFTGFTVGRGKTLVLAISPVIGLSAAIMAKILS 237
QY 240 SFTDKEILAYAKAGAAVEEVLAIIRTVIAFGQKKELERYKNLEBAKIGICAKITANI 299
DB 238 SFTDKEILAYAKAGAAVEEVLAIIRTVIAFGQKKELERYKNLEBAKIGICAKITANI 297
QY 300 SIGAFLLIYASTALAFWGTSLVLSSEYSGQVLTVEFSVLICAFSIGQSPSIEAFAN 359
DB 298 SIGAFLLIYASTALAFWGTSLVLSSEYSGQVLTVEFSVLICAFSIGQSPSIEAFAN 357
QY 360 ARGAAYEIRKIIDNKPSIDISYSGSKHPDNKGNLEFNHVSYSRKEVAILKGLNLKY 419
DB 358 ARGAAYEIRKIIDNKPSIDISYSGSKHPDNKGNLEFNHVSYSRKEVAILKGLNLKY 417
QY 420 QSGQTVALVNSGCGKSTTVQMLQRLYPTDGMVCIDGODITITNVRRLREITVSGEP 479
DB 418 QSGQTVALVNSGCGKSTTVQMLQRLYPTDGMVCIDGODITITNVRRLREITVSGEP 477
QY 480 VLFATYIAENIRGRENVMTDETEKAVKANAYDIPMKLPNKFDTLVGERGAQLSGGQ 539
DB 478 VLFATYIAENIRGRENVMTDETEKAVKANAYDIPMKLPNKFDTLVGERGAQLSGGQ 537
QY 540 RIATARALVNPRIILLDEATSALDTESEAVVQVALDKARKGRTTIVIAHRLSTVRNDV 599
DB 538 RIATARALVNPRIILLDEATSALDTESEAVVQVALDKARKGRTTIVIAHRLSTVRNDV 597
QY 600 IAGFDGVIYKGNHDELMKEKGIYFKLYTMOTRGNEIELENADESKSESDALEMSPKD 659
DB 598 IAGFDGVIYKGNHDELMKEKGIYFKLYTMOTRGNEIELENADESKSESDALEMSPKD 657
QY 660 SGSSLIRKRSRHSIHAPOGQDRKLTKEJLNENPVSPVFWRLIKLNSTEMPVYVIGFC 719
DB 658 SGSSLIRKRSRHSIHAPOGQDRKLTKEJLNENPVSPVFWRLIKLNSTEMPVYVIGFC 717
QY 720 AIINGGLQAPAFSIIFSRILIGITRDEDPETKRONSMFVLPVIGITISFTTFQGTFF 779
DB 718 AIINGGLQAPAFSIIFSRILIGITRDEDPETKRONSMFVLPVIGITISFTTFQGTFF 777
QY 780 GKAEIILTKRLRYVRFVSMRLRODYSWPDDBKNTGALTTRLANDAAQVGAIGSRILAVIT 839
DB 778 GKAEIILTKRLRYVRFVSMRLRODYSWPDDBKNTGALTTRLANDAAQVGAIGSRILAVIT 837
QY 840 QNINLNGIGIISITLYCMQTLTLLALVPIITAIAGVEMKMLSGQALDKKKELEGAGKIA 899
DB 838 QNINLNGIGIISITLYCMQTLTLLALVPIITAIAGVEMKMLSGQALDKKKELEGAGKIA 897
QY 900 TEATENFTVSVLTREQKFEYMAQSLQVPYRNSLKAHIFGVSESIQAMMYFVAGCF 959
DB 898 TEATENFTVSVLTREQKFEYMAQSLQVPYRNSLKAHIFGVSESIQAMMYFVAGCF 957
QY 960 RFGATLVANEMNQDVLVFSALVFCAMVAVGVSSPAPDYAKAKVSAHVIMIEKSP 1019

```

DB 958 RFAGYLAHAKLMSEFEDLVFSAIVFGAMAVGVSSFAADYAKAKISAAHITMITETPL 1017
QY 1020 IDSTSPHGLKPNLEGNVFNENEVYPTRPDIPVLQGLSLEVKKGQTLALVSSGCGKS 1079
DB 1018 IDSTSTGLKPNLEGNVFNENEVYPTRPDIPVLQGLSLEVKKGQTLALVSSGCGKS 1077
QY 1080 TVVOLLERFYDPLAGSVLIDGKEIKHLANTQWMLRAHGLGYSOBPIIFDCSIAENTAYGNS 1139
DB 1078 TVVOLLERFYDPLAGSVLIDGKEIKHLANTQWMLRAHGLGYSOBPIIFDCSIAENTAYGNS 1137
QY 1140 RVVSHHEIMQAKEANIHIFIEFLPEKYNTRVGDGKTOLSGGOKORAIARALYRQPHIL 1199
DB 1138 RVVSEIIVRAAKAENIHAFIESLPMKYSTKVGDKGTOLSGGOKORAIARALYRQPHIL 1197
QY 1200 LIDDEATSAIDTESEKVVQALDKARGRCIYIAHRLSTIQNADLIYVFQNGKYKEHGT 1259
DB 1198 LIDDEATSAIDTESEKVVQALDKARGRCIYIAHRLSTIQNADLIYVFQNGKYKEHGT 1257
QY 1260 QQLLAKGIFYSMVSYOAGAKR 1281
DB 1258 QQLLAKGIFYSMVSYOAGAKR 1279

RESULT 12
AAE18967
ID AAE18967 standard; Protein; 1280 AA.
AC AAE18967;
DT 21-MAY-2002 (first entry)
DE Human P-gp mutant, MDR1-L56P (8-A4).
KM Human; P-glycoprotein; vaccine; tumour; chemotherapeutic agent; cancer;
KN haematopoietic cell; cytotoxicity; cytostatic; P-gp; mutant; mutein.
OS Homo sapiens.
OS Synthetic.
FH Key Location/Qualifiers
FT Misc-difference 56 /note="Wild type Leu substituted with Pro"
PN WO200210205-A2.
PD 07-FEB-2002.
PE 01-AUG-2001; 2001WO-US24560.
PR 01-AUG-2000; 2000US-222313P.
PA (UNII ) UNIV ILLINOIS FOUND.
XX Ruth A, Roninson I;
XX PI
XX DR WPI; 2002-206182/26.
XX
PT New human P-glycoprotein mutants for treating cancer, comprises an
PT ability to confer increased resistance to chemotherapeutic drugs
PT relative to wild type P-glycoprotein or P-glycoprotein having a glycine
PT to valine substitution
XX
XX Claim 42; Fig 3; 47pp; English.
XX
CC The present invention relates to an isolated mutant human P-glycoprotein
CC molecule. The invention is used as vaccine. An antibody specific for the
CC mutant P-glycoprotein is useful for determining whether a human tumour
CC sample comprises tumour cells expressing mutant human P-glycoprotein. A
CC recombinant expression construct comprising nucleic acid encoding mutant
CC P-glycoprotein is useful for administering an increased amount of a
CC chemotherapeutic agent to an individual with cancer, by transducing ex
CC vivo haematopoietic cells from the individual with the recombinant

```

```

CC expression construct, reintroducing the transduced haematopoietic cells
CC into the individual and administering an increased amount of a
CC chemotherapeutic agent without consequent haematopoietic cytotoxicity.
CC The mutant P-glycoprotein is useful for treating and diagnosing cancer
CC and for screening compounds for the capacity to disrupt binding of
CC cytotoxic drugs with the mutant P-glycoprotein or to disrupt drug efflux
CC from cells expressing the mutant P-glycoprotein. The present sequence is
CC human P-glycoprotein (P-gp) mutant, MDR1-L56P (8-A4)
CC (multidrug resistance).
XX
SQ Sequence 1280 AA;
Query Match 90.5%; Score 5859; DB 23; Length 1280;
Best Local Similarity 90.7%; Pred. No. 0;
Matches 1163; Conservative 54; Mismatches 61; Indels 4; Gaps 4;
QY 1 MPEGGRKSA-EKNPMKKGSKSKNEKEKKPYSTRAMFRYSNMDRLYMGVMAAI 59
DB 1 MDEGDRNGAKKKKPFKLNKSEK-DKEKKPYVSFMSFRYSNMWLDKLYMVGVFAAI 59
QY 60 IHGAALPLMMVLVEGNMTDSFANAGISRKTPPVIIINESITNNTQHPINHLEEMTAYY 119
DB 60 IHGAGLPMLMLVEGEMTDIFANAG-NLEDLMSNTNNSDINDTGFEMK-LEEDMTAYY 117
QY 120 YSGIGAGVLAAYIOVSWCILAQRQILIKRQFFHAIMRQELGMPDVHVGELNRLTD 179
DB 118 YSGIGAGVLAAYIOVSWCILAQRQILIKRQFFHAIMRQELGMPDVHVGELNRLTD 177
QY 180 DYSKINIEGDKIGMFHFFTFGTFVGTGRGKTLVILASPLVGLSAIAIMAKIIS 239
DB 178 DYSKINIEGDKIGMFHFFTFGTFVGTGRGKTLVILASPLVGLSAIAIMAKIIS 237
QY 240 SFTDKELLAYAKAGAAVEVLAIRTVIAFGQKKELEERYKNLEAKGIGIKATANI 299
DB 238 SFTDKELLAYAKAGAAVEVLAIRTVIAFGQKKELEERYKNLEAKGIGIKATANI 297
QY 300 SIGAAFLITVASYALAFMYGTSVLSESYIGOVLYVFESVILGASISQASPIAFAN 359
DB 298 SIGAAFLITVASYALAFMYGTSVLSESYIGOVLYVFESVILGASISQASPIAFAN 357
QY 360 ARGAAVEIFKLIIDNKPISIDSYSGHKNPDINKNLEKKNVHFSYPSRKEKILKGLNLV 419
DB 358 ARGAAVEIFKLIIDNKPISIDSYSGHKNPDINKNLEKKNVHFSYPSRKEKILKGLNLV 417
QY 420 QSGOTVALVSGSGKSTVQVMQRLYDPDGAVCIDGODITINVRHLREITGVVSGEP 479
DB 418 QSGOTVALVSGSGKSTVQVMQRLYDPDGAVCIDGODITINVRHLREITGVVSGEP 477
QY 480 VLEATTIAENIRYGRENVIMDEIEKAVKEANAYDFIMKLPKFDVLVGRGAQLSGGOK 539
DB 478 VLEATTIAENIRYGRENVIMDEIEKAVKEANAYDFIMKLPKFDVLVGRGAQLSGGOK 537
QY 540 RIAIARALVNRKILLIDDEATSAIDTESAAVVOVALDKARKGRTIYIAHRLSTVRNDV 599
DB 538 RIAIARALVNRKILLIDDEATSAIDTESAAVVOVALDKARKGRTIYIAHRLSTVRNDV 597
QY 600 IAGFDGVIVKGNHDELKKEKGYFKLYTMOTRNEIELENAESESSEDALEMSXD 659
DB 598 IAGFDGVIVKGNHDELKKEKGYFKLYTMOTRNEIELENAESESSEDALEMSXD 657
QY 660 SSGSLIKRSTRSRSHAPOGDRLGTGEDLENVPVPSFIRLKLNSTEMPVYVGC 719
DB 658 SRSLSIKRSTRSRVSGAODRKLTSEALDESIPVPSFIRKLNLTNEMPVYVGC 717
QY 720 AIINGLOPAFLISRIIGITRDEDEPTKRONSMSVLFVLGITSITFFLOGFTF 779
DB 718 AIINGLOPAFLISRIIGITRDEDEPTKRONSMSVLFVLGITSITFFLOGFTF 777
QY 780 GRAGETLKRLLRYVFRSKLRQDVSWFDDPKNTTGALTTRLANDAAOVKAGISRLAVIT 839
DB 778 GRAGETLKRLLRYVFRSKLRQDVSWFDDPKNTTGALTTRLANDAAOVKAGISRLAVIT 837
QY 840 QMIANLNGIITISLTYQMQLTLLLAIVPIITAIAGVVEKMLSGQALDKRKELEGAKIA 899

```

|||||
Db 838 QNINNLGIIISFIYQMQLTLLLAIVPIIATAGVEKMLSGQALKKKELEGAGKIA 897
QY 900 TEALENFTVYSLIREQEFEMYNOSLOVYRNSLRKAHIFGVSESITQAMMFYAGCF 959
Db 898 TEALENFTVYSLIREQEFEMYNOSLOVYRNSLRKAHIFGVSESITQAMMFYAGCF 957
QY 960 RFGALVANEEMFNQDVLVFSALVFGAMAVGVSSFPADYAKAKYSAAHVIMIEKSPIL 1019
Db 958 RFGALVANEEMFNQDVLVFSALVFGAMAVGVSSFPADYAKAKYSAAHVIMIEKSPIL 1017
QY 1020 IDSVPBHLKNTLEGANTFENEVFNYPTRPDIPVLQGLSLEVKKGQTLALVSSSCGKS 1079
Db 1018 IDSVPBHLKNTLEGANTFENEVFNYPTRPDIPVLQGLSLEVKKGQTLALVSSSCGKS 1077
QY 1080 TVVOLLERYPDLAGSVLIDKEIKHLNVQMLRAHLGIVSEPLTFPCSTIAENIAYGDS 1139
Db 1078 TVVOLLERYPDLAGSVLIDKEIKHLNVQMLRAHLGIVSEPLTFPCSTIAENIAYGDS 1137
QY 1140 RVVSHBEIQAQKAEINHHFETLPEKYNTRYGDKGTQLSGGQKQRTAIAIALVROPHIL 1199
Db 1138 RVVSHBEIQAQKAEINHHFETLPEKYNTRYGDKGTQLSGGQKQRTAIAIALVROPHIL 1197
QY 1200 LIDEATSALDTESEKVVQVQALDKAREGRTCIYIAHRLSTIONADLIYVFQNGKYEKGT 1259
Db 1198 LIDEATSALDTESEKVVQVQALDKAREGRTCIYIAHRLSTIONADLIYVFQNGKYEKGT 1257
QY 1260 QQLLAKGIFYFMSVQAGAKR 1281
Db 1258 QQLLAKGIFYFMSVQAGAKR 1279

RESULT 13
AAB81064
ID AAB81064 standard; Protein; 1280 AA.
XX AAB81064;
AC AAB81064;
XX 25-JUN-2001 (first entry)
DT 25-JUN-2001 (first entry)
XX Cynomologous monkey P-glycoprotein variant 1.
DE Cynomologous monkey P-glycoprotein; Pgp; multidrug transporter; MDR1;
XX Cynomologous monkey; P-glycoprotein; Pgp; multidrug transporter; MDR1;
KW efflux pump.
XX Macaca fascicularis.
OS Macaca fascicularis.
XX WO200123565-A1.
PN WO200123565-A1.
XX 05-APR-2001.
PD 05-APR-2001.
XX 28-SEP-2000; 2000WO-US26592.
PF 28-SEP-2000; 2000WO-US26592.
XX 28-SEP-1999; 99US-0156921.
PR 12-OCT-1999; 99US-0156921.
XX 12-OCT-1999; 99US-0156921.
PA (GENT-) GENTEST CORP.
XX (GENT-) GENTEST CORP.
PI Stocker PJ, Steimel-Crespi DT, Crespi CL;
XX Stocker PJ, Steimel-Crespi DT, Crespi CL;
XX WPI: 2001-316136/33.
DR N-PSDB: AAF86127.
XX N-PSDB: AAF86127.
XX Novel isolated nucleic acid encoding cynomologous monkey P-glycoprotein
PT (Pgp) and homologous Pgp polypeptides are useful for predicting
PT bioavailability of compound and increasing Pgp transporter activity in
PT cell
XX
PS Claim 9; Page 57-59; 84pp; English.
XX This invention relates to a polynucleotide sequence encoding a
CC cynomologous monkey P-glycoprotein (Pgp), and an allelic variant of the
CC Pgp protein. Pgp, also known as multidrug transporter, MDR1 is a member

CC of the ABC transporter superfamily. The enzyme serves as an efflux pump
CC exporting small molecules across the cell membrane. The invention
CC includes a cynomologous monkey (Macaca fascicularis) Pgp coding sequence
CC and protein, and also that of an allelic variant. The Pgp polynucleotide
CC sequence is useful for increasing Pgp transporter activity in a cell.
CC Antisense sequences of the cDNA are useful for inhibiting Pgp transport
CC activity in a mammalian cell. They may also be used for increasing the
CC bioavailability of a drug. The present sequence represents the
CC cynomologous monkey P-glycoprotein.
XX
XX Sequence 1280 AA:
Query Match 90.4%; Score 5857; DB 22; Length 1280;
Best Local Similarity 90.6%; Pred. No. 0;
Matches 1162; Conservative 59; Mismatches 57; Indels 4; Gaps 4;
QY 1 MDPEGGRKSAE-KNFWKMGKSKKNEKKEKPPVSTFAFRRSNMLDRLYMVGMAI 59
Db 1 MDPEGGRKSAE-KNFWKMGKSKKNEKKEKPPVSTFAFRRSNMLDRLYMVGMAI 59
QY 60 IHGALPLMLVFGMDTDFANAG-NIGDGLALLTNSNTIDVYVYN-LEEDMTYAY 117
Db 60 IHGALPLMLVFGMDTDFANAG-NIGDGLALLTNSNTIDVYVYN-LEEDMTYAY 117
QY 120 YSGIGAGVLAAYIGVSPWCLAGROILKIRKOFFHAIMROEIGMFPVHVGELNRLTD 179
Db 118 YSGIGAGVLAAYIGVSPWCLAGROILKIRKOFFHAIMROEIGMFPVHVGELNRLTD 177
QY 180 DYSKINEGIGDKIGMFFSHATFETGFTVGTFRCKMLTLVLAISPLVIGSAAMAKILS 239
Db 178 DYSKINEGIGDKIGMFFSHATFETGFTVGTFRCKMLTLVLAISPLVIGSAAMAKILS 237
QY 240 SFTDKELAVAKAGVAEEVLAIRTVIAFGQKKELEKRYKNELEAKGIGIKKAITANI 299
Db 238 SFTDKELAVAKAGVAEEVLAIRTVIAFGQKKELEKRYKNELEAKGIGIKKAITANI 297
QY 300 SIGAFLIYASALAFWGTGLVLSSEYSGIVTFVFSVLIGAFSGQSPSIEAFAN 359
Db 298 SIGAFLIYASALAFWGTGLVLSSEYSGIVTFVFSVLIGAFSGQSPSIEAFAN 357
QY 360 ARGAAYEFKIIDNKPSIDSYSGHGPDKIKGNLEFNHVSYPKREYVILKGLNKV 419
Db 358 ARGAAYEFKIIDNKPSIDSYSGHGPDKIKGNLEFNHVSYPKREYVILKGLNKV 417
QY 420 QSGQVAVLNGSCGKSTVQMLQRLYDPTDGMVQIDODRTIIVRLRETGVVSGEP 479
Db 418 QSGQVAVLNGSCGKSTVQMLQRLYDPTDGMVQIDODRTIIVRLRETGVVSGEP 477
QY 480 VLFATTTAENIRYGRNVMDIEKAVKANAYDFIMKLPQKFDPLVGERGQSLSGGQK 539
Db 478 VLFATTTAENIRYGRNVMDIEKAVKANAYDFIMKLPQKFDPLVGERGQSLSGGQK 537
QY 540 RIATARALVBNPKILLDEATSALDTESEAVVOVALDKARKRTTIVIAHRLSTYRNADV 599
Db 538 RIATARALVBNPKILLDEATSALDTESEAVVOVALDKARKRTTIVIAHRLSTYRNADV 597
QY 600 IAGPDDGVYIEKGNHDELMKEGITYFLVYMTGRNEIELENATGSEKESDALBMSPD 659
Db 598 IAGPDDGVYIEKGNHDELMKEGITYFLVYMTGRNEIELENATGSEKESDALBMSPD 657
QY 660 SSSSLIKRSTRSRIRAHPOGODRKLGTEDLNENPVVSWRLKLNSTEWYFVVGJFC 719
Db 658 SSSSLIKRSTRSRIRAHPOGODRKLGTEDLNENPVVSWRLKLNSTEWYFVVGJFC 717
QY 720 AINGGLOPAFSTIFSKRTIGFTDEDEPERKRONSNWFLVFLVGLISITFFLOGFTF 779
Db 718 AINGGLOPAFSTIFSKRTIGFTDEDEPERKRONSNWFLVFLVGLISITFFLOGFTF 777
QY 780 GRAGEITLKRRLRYWFRSMLRODVSWFDDPKNTGALTTLRLANDAQVKAISRLAIVT 839
Db 778 GRAGEITLKRRLRYWFRSMLRODVSWFDDPKNTGALTTLRLANDAQVKAISRLAIVT 837
QY 840 QNINNLGIIISFIYQMQLTLLLAIVPIIATAGVEKMLSGQALKKKELEGAGKIA 899

```
Db 838 QNIAHMGITIIISLIYGMOLTTLLAIYIIAIGAVEMKMLSGALDKKELSEGAGIA 897
QY 900 TEAIEHFRFVSLTRPQKFEVMAAOSIQVPYNSLRKAHIGVSTIQAMMTSYACF 959
Db 898 TEAIEHFRFVSLTRPQKFEVMAAOSIQVPYNSLRKAHIGVSTIQAMMTSYACF 957
QY 960 REGAVIVANEFNFODVLVFSIAVFGAMAVGVSSFPADAKAKVSAHYIMIEKSP 1019
Db 958 REGAVIVANEFNFODVLVFSIAVFGAMAVGVSSFPADAKAKVSAHYIMIEKSP 1017
QY 1020 IDSYPHGAKPTLBSGNTFNFNVTNPTRPDIYVGLSLSEVKGQTLALVSGSGGKS 1079
Db 1018 IDSYPHGAKPTLBSGNTFNFNVTNPTRPDIYVGLSLSEVKGQTLALVSGSGGKS 1077
QY 1080 TVVOLLERYDPLAGSVLIDGKEIKHLVWOMLRHILGIVSOEPLTFQCSIAENIAYGDS 1139
Db 1078 TVVOLLERYDPLAGSVLIDGKEIKHLVWOMLRHILGIVSOEPLTFQCSIAENIAYGDS 1137
QY 1140 RVVSHETMQAAKEANIHFIETLPEKYNTRVGDGTQLSGGQKQRTAIRALVROPHIL 1199
Db 1138 RVVSHETMQAAKEANIHFIETLPEKYNTRVGDGTQLSGGQKQRTAIRALVROPHIL 1197
QY 1200 ILDEKTSALDPESEKVOEALDKAREGTCIVIAHRLSTIONADLIYVFGNGVKEHGT 1259
Db 1198 ILDEKTSALDPESEKVOEALDKAREGTCIVIAHRLSTIONADLIYVFGNGVKEHGT 1257
QY 1260 QQLLAQKGIYFSMVSVOGAKR 1281
Db 1258 QQLLAQKGIYFSMVSVOGAKR 1279

RESULT 14
AAE18968
ID AAE18968 standard; Protein: 1280 AA.
AC AAE18968;
XX
XX 21-MAY-2002 (first entry)
XX
XX Human P-gp mutant, MDR1-1190F/M197F (5-C3).
XX
XX Human; P-glycoprotein; vaccine; tumour; chemotherapeutic agent; cancer;
XX haematopoietic cell; cytotoxicity; cytostatic; P-gp; mutant; mutleln.
XX
XX Homo sapiens.
XX
XX Synthetic.
XX
XX Key Location/Qualifiers
XX MISC-difference 190
XX MISC-difference 197 /note= "Wild type Ile substituted with Phe"
XX MISC-difference 197 /note= "Wild type Met substituted with Phe"
XX
XX W0200210205-A2.
XX
XX 07-FEB-2002.
XX
XX 01-AUG-2001: 2001WO-US24560.
XX
XX 01-AUG-2000: 2000GS-222313P.
XX
XX (UNIL ) UNIL ILLINOIS FOUND.
XX
XX Ruth A, Roninson I;
XX
XX WPI; 2002-206182/26.
XX
XX New human P-glycoprotein mutants for treating cancer. comprises an
XX ability to confer increased resistance to chemotherapeutic drugs
XX relative to wild type P-glycoprotein or P-glycoprotein having a glycine
XX to valine substitution
```

```
PS Claim 42: Fig 3; 47bp; English.
XX
XX The present invention relates to an isolated mutant human P-glycoprotein
XX molecule. The invention is used as vaccine. An antibody specific for the
XX mutant P-glycoprotein is useful for determining whether a human tumour
XX sample comprises tumour cells expressing mutant human P-glycoprotein. A
XX recombinant expression construct comprising nucleic acid encoding mutant
XX P-glycoprotein is useful for administering an increased amount of a
XX chemotherapeutic agent to an individual with cancer, by transducing ex
XX vivo haematopoietic cells from the individual with the recombinant
XX expression construct, reintroducing the transduced haematopoietic cells
XX into the individual and administering an increased amount of a
XX chemotherapeutic agent without consequent haematopoietic cytotoxicity.
XX The mutant P-glycoprotein is useful for treating and diagnosing cancer
XX and for screening compounds for the capacity to disrupt binding of
XX cytotoxic drugs with the mutant P-glycoprotein or to disrupt drug efflux
XX from cells expressing the mutant P-glycoprotein. The present sequence is
XX human P-glycoprotein (P-gp) mutant, MDR1-1190F/M197F (5-C3)
XX
XX Sequence 1280 AA;
SQ
Query Match 90.4%; Score 5856; DB 23; Length 1280;
Best Local Similarity 90.6%; Pred. No. 0;
Matches 1162; Conservative 54; Mismatches 62; Indels 4; Gaps 4;
QY 1 MDPEGKRGSA-EKNFMKMGKSKKNEKPPVSTFAMFRYSNMIDRLYMLVGTMAI 59
Db 1 MDLEDGKRGSAKKNFELNKKSEK-DKEKKPVSYSFMSFRYSNMIDRLYMLVGTMAI 59
QY 60 IHGAALPLMLVFGNMTDPSFANAGISRNKTPVYINSITNNQTHLNLHEEMTYYAY 119
Db 60 IHGAALPLMLVFGNMTDPSFANAGISRNKTPVYINSITNNQTHLNLHEEMTYYAY 117
QY 120 YSIGAGVLAAYIQVSWFCIAAGRQILIRKOFHAIMROEIGWPDVHDVGLNRLTD 179
Db 118 YSIGAGVLAAYIQVSWFCIAAGRQILIRKOFHAIMROEIGWPDVHDVGLNRLTD 177
QY 180 DVSKINGIGDKIGMFHSHATFTFTGIVGFTRGKILTVIIAISPYGLISAIAIAKILIS 239
Db 178 DVSKINGIGDKIGMFHSHATFTFTGIVGFTRGKILTVIIAISPYGLISAIAIAKILIS 237
QY 240 SFTDKELLAYAKAGAAVEEYLAIRFYVIAAGGOKKLEBYNNLEAKSIGIKKAITANI 299
Db 238 SFTDKELLAYAKAGAAVEEYLAIRFYVIAAGGOKKLEBYNNLEAKSIGIKKAITANI 297
QY 300 SIGAFLIITASYALAFWGTSLVLSSEYSIGVLTVPFSVLGAFSIGQASPIEAFAN 359
Db 298 SIGAFLIITASYALAFWGTSLVLSSEYSIGVLTVPFSVLGAFSIGQASPIEAFAN 357
QY 360 ARGAAVEIFKIIDNKSISYSKSGKPNINGNLEFKNVHFSYSRKEVYIKGLNLKV 419
Db 358 ARGAAVEIFKIIDNKSISYSKSGKPNINGNLEFKNVHFSYSRKEVYIKGLNLKV 417
QY 420 QSGQVAVLWNSGCGKSTTVQMLQRLYDPTDGMVCIDGODITINVRHLREITGVVSE 479
Db 418 QSGQVAVLWNSGCGKSTTVQMLQRLYDPTDGMVCIDGODITINVRHLREITGVVSE 477
QY 480 VLFATITAEIRYGRNVTMDLEKAVKANAYDITMKLPKFDTLVGERGAQLSGGOK 539
Db 478 VLFATITAEIRYGRNVTMDLEKAVKANAYDITMKLPKFDTLVGERGAQLSGGOK 537
QY 540 RIAIARALVBNPKIILLDEATSAIDTESAAVVOVALDKARKGRITIVIAHLSVIRNADV 599
Db 538 RIAIARALVBNPKIILLDEATSAIDTESAAVVOVALDKARKGRITIVIAHLSVIRNADV 597
QY 600 IAGFDGVIVGNDHDELMEKEGIYFKLVMTQGRNLELEENATGESKSESDALEMSPKD 659
Db 598 IAGFDGVIVGNDHDELMEKEGIYFKLVMTQGRNLELEENATGESKSESDALEMSPKD 657
QY 660 SSSSLIKRSTRSIIHAPGODRKIGTKEDLNENVPVPSFRILKLSTEMPPYVYVIFC 719
Db 658 SSSSLIKRSTRSIIHAPGODRKIGTKEDLNENVPVPSFRILKLSTEMPPYVYVIFC 717
```

```
QY 720 AIIINGIOPAFSIIFRSIIIGTIFTRDEDEPETKRONSNMFSVLVLGIITSEFTPELQGFTE 779
    |||||
DB 718 AIIINGIOPAFSIIFRSIIIGTIFTRDEDEPETKRONSNMFSVLVLGIITSEFTPELQGFTE 777
QY 780 GRAGEIILTRKLRMYFRSMRLQDVSMWEDDPKNTGALTTRLANDAAQYKAGISRLAVIT 839
    |||||
DB 778 GRAGEIILTRKLRMYFRSMRLQDVSMWEDDPKNTGALTTRLANDAAQYKAGISRLAVIT 837
QY 840 QNANIGTIIISLIYGMQLTLLILAIPIIATAGVEMKMLSGOALKDKKELEGACKIA 899
    |||||
DB 838 QNANIGTIIISLIYGMQLTLLILAIPIIATAGVEMKMLSGOALKDKKELEGACKIA 897
QY 900 TEAIEFRIVVSLTRQKREXYMYAOSLOYPRNSLRKAHIFGSFSTIQAMMTSYACGF 959
    |||||
DB 898 TEAIEFRIVVSLTRQKREXYMYAOSLOYPRNSLRKAHIFGSFSTIQAMMTSYACGF 957
QY 960 REGAYIVANEPMFQOVLLVFSNAIVFGAMAVGVSSPAPDYAKAKVSAAHYIMITEKSPL 1019
    |||||
DB 958 REGAYIVANEPMFQOVLLVFSNAIVFGAMAVGVSSPAPDYAKAKVSAAHYIMITEKSPL 1017
QY 1020 IDSXSPHGLKPNTEGNTFENEVFNYPTRPDIPVLOGLSLEVKKGQTLALVSSGCGKS 1079
    |||||
DB 1018 IDSXSPHGLKPNTEGNTFENEVFNYPTRPDIPVLOGLSLEVKKGQTLALVSSGCGKS 1077
QY 1080 TVVQLLERFYDPLAGSVLIDGKEIKHLNQWLRKHLGIYSOEPILFDCSIAENIAYGDS 1139
    |||||
DB 1078 TVVQLLERFYDPLAGSVLIDGKEIKHLNQWLRKHLGIYSOEPILFDCSIAENIAYGDS 1137
QY 1140 RYVSHIEIMQAKENANIHFIETLPEKYNTRYGDKGTQLSGQOKRIATARALVROPHIL 1199
    |||||
DB 1138 RYVSHIEIMQAKENANIHFIETLPEKYNTRYGDKGTQLSGQOKRIATARALVROPHIL 1197
QY 1200 LLDDEATSDTESSEKVVQALDKAREGRICIVIAHFLSTIQNADLIYVQNKVHEHGH 1259
    |||||
DB 1198 LLDDEATSDTESSEKVVQALDKAREGRICIVIAHFLSTIQNADLIYVQNKVHEHGH 1257
QY 1260 QQLLAQKGIYFSMYSVOAGAKR 1281
    |||||
DB 1258 QQLLAQKGIYFSMYSVOAGAKR 1279

RESULT 15
AAB81065
ID AAB81065 standard; Protein; 1283 AA.
XX
AC AAB81065;
XX
DT 25-JUN-2001 (first entry)
XX
DE Cynomologous monkey P-glycoprotein variant 2.
XX
KW Cynomologous monkey; P-glycoprotein; PGP; multidrug transporter; MDRL;
    efflux pump.
XX
OS Macaca fascicularis.
XX
FH Key location/Qualifiers
FT Misc-difference 93..95 "An additional 3 amino acids are present compared
    to PGP variant AAB81064"
FT
FT
FT
PN W0200123565-A1.
PD
PD 05-APR-2001.
PF 28-SEP-2000; 2000WO-US26592.
PR 28-SEP-1999; 99US-0156921.
PR 12-OCT-1999; 99US-0158818.
XX
XX (GENE-) GENTEST CORP.
XX
PA
```

```
PI Stocker PJ, Steimel-Crespi DT, Crespi CL.
XX WPI: 2001-316136/33.
DR N-PSDB; AAF86128.
XX
PT Novel isolated nucleic acid encoding cynomologous monkey P-glycoprotein
    (PGP) and homologous PGP polypeptides are useful for predicting
    bioavailability of compound and increasing PGP transporter activity in
    cell.
XX
XX Claim 9, Page 65-68; 84pp; English.
XX
CC This invention relates to a polynucleotide sequence encoding a
    cynomologous monkey P-glycoprotein (PGP), and an allelic variant of the
    PGP protein. PGP, also known as multidrug transporter, MDRL is a member
    of the ABC transporter superfamily. The enzyme serves as an efflux pump
    exporting small molecules across the cell membrane. The invention
    includes a cynomologous monkey (macaca fascicularis) PGP coding sequence
    and protein, and also that of an allelic variant. The PGP polynucleotide
    sequence is useful for increasing PGP transporter activity in a cell.
    Antisense sequences of the cDNA are useful for inhibiting PGP transport
    activity in a mammalian cell. They may also be used for increasing the
    bioavailability of a drug. The present sequence represents the
    cynomologous monkey P-glycoprotein variant 2. The protein has an
    additional 3 amino acids when compared to PGP variant 1 (AAB81065).
XX
XX Sequence 1283 AA;
XX
Query Match 90.4%; Score 5854.5; DB 22; Length 1283;
Best Local Similarity 90.3%; Pred. NO. 0;
Matches 1164; Conservative 57; Mismatches 53; Indels 15; Gaps 5;

QY 1 MDEEGGRKGSAAE-KNFMKMKSKSKNEKKEKPPYVSTAMFYSWMDRLVMTVGTMAI 59
DB 1 MDEEGGRKGSAAE-KNFMKMKSKSKNEKKEKPPYVSTAMFYSWMDRLVMTVGTMAI 59
QY 60 IHGAGLPIMLVGNMTPSPANG-----ISRKTPFVYINESITNTOFHINLREE 112
DB 60 IHGAGLPIMLVGNMTPSPANG-----ISRKTPFVYINESITNTOFHINLREE 112
QY 60 IHGAGLPIMLVGNMTPSPANG-----ISRKTPFVYINESITNTOFHINLREE 112
DB 60 IHGAGLPIMLVGNMTPSPANG-----ISRKTPFVYINESITNTOFHINLREE 112
QY 113 MTTVAAYVSGIGAVLVAAVYQVSEFWCLAAGRQILKIRKQFHAIMROEIGFVDHVDGE 172
DB 114 MTTVAAYVSGIGAVLVAAVYQVSEFWCLAAGRQILKIRKQFHAIMROEIGFVDHVDGE 173
QY 173 LNTPLTDVSKINEGIDKIGMFESHATFETGFIYGTGKMLTVLAIASPVGLSAA 232
DB 174 LNTPLTDVSKINEGIDKIGMFESHATFETGFIYGTGKMLTVLAIASPVGLSAA 233
QY 233 IWAKIISFTDKELLATAKAGAAVEEVLAIIRTVIAFGQKKELERYNKNLEBAKGIGIK 292
DB 234 IWAKIISFTDKELLATAKAGAAVEEVLAIIRTVIAFGQKKELERYNKNLEBAKGIGIK 293
QY 293 KAITANISIGAAFLITIASTALAFWYGTSLVSESEVSIQOVLPFVSYLIGAFSIGOASP 352
DB 294 KAITANISIGAAFLITIASTALAFWYGTSLVSESEVSIQOVLPFVSYLIGAFSIGOASP 353
QY 353 STEAPANRGAAYEFETIDNKPISIDYSKSGKHPNIKGNLEFRKNVHSYSPRKEVKIL 412
DB 354 STEAPANRGAAYEFETIDNKPISIDYSKSGKHPNIKGNLEFRKNVHSYSPRKEVKIL 413
QY 413 KGLNLKVOGQTVALLVNGSGGKSTVOLMQLYDPTDGNVICIDGODIRTNVRHLRETT 472
DB 414 KGLNLKVOGQTVALLVNGSGGKSTVOLMQLYDPTDGNVICIDGODIRTNVRHLRETT 473
QY 473 GVSQEPVLFATTAENIRYGRVNTYDELEKAVKENAADYIMKLPNKFDTLVGERGAQ 532
DB 474 GVSQEPVLFATTAENIRYGRVNTYDELEKAVKENAADYIMKLPNKFDTLVGERGAQ 533
QY 533 IISGOKORITATARALVRNPKIILDEATSAIDTESAVVOVALDKARKGRTTIVIAHRLS 592
DB 534 IISGOKORITATARALVRNPKIILDEATSAIDTESAVVOVALDKARKGRTTIVIAHRLS 593
QY 593 TVRNADVIAAGDDGVYVERGNHDELMKEKGIYFLVTMOTRGNEIELENAITGESKESDA 652
```



```

Db 594 TVRADVLAGPDGIVLVEGNHDELMEKGIYERKLYTMQTAGNELEENADESKSEIDT 653
QY 653 LEMSPKDSGSSSLIKRRSTRSRSHAFQODRKLTGREDLENENPVPSFWRLKLNSTENPY 712
Db 654 LEMSHDSGSSSLIKRRSTRSRVSGQDRKLTKEALDESIPVPSFWRMKLNTEMPY 713
QY 713 FVVGICFALINGLOPAFSITPSRIIGFTREDDEPETKRONSMFSVLTVIGIISFTTF 772
Db 714 FVVGVCALINGLOPAFAVIFSRTIGFTRNDAETKRONSMFSILFLVIGIYSFTTF 773
QY 773 FLQGFTEGAGELITKRLRYWFRSMRQDVSWPDPKNTGALTTRLANDAAYKGAIG 832
Db 774 FLQGFTEGAGELITKRLRYWFRSMRQDVSWPDPKNTGALTTRLANDAAYKGAIG 833
QY 833 SRLAVITOMIANLGTGIIISLIYGWQTLTLLAIVPIAIAAGVEMKMLSGQALDKKEL 892
Db 834 SRLAITOMIANLGTGIIISLIYGWQTLTLLAIVPIAIAAGVEMKMLSGQALDKKEL 893
QY 893 EGAGKLTATEIENFRTVVSILPREOKFEYMAQSLQVPRNSLRKAHIGVSPSTQAMMY 952
Db 894 EGAGKLTATEIENFRTVVSILPREOKFEYMAQSLQVPRNSLRKAHIGVSPSTQAMMY 953
QY 953 FSYAGCFRGAYLVANEFMNEFODVLVFSALVEGAMAVGQSSPADYAKAKVSAHYIM 1012
Db 954 FSYAGCFRGAYLVANEFMNEFODVLVFSALVEGAMAVGQSSPADYAKAKVSAHYIM 1013
QY 1013 IIEKSPILDSYSPHGLKPTLEGNTFENEVENYPTRPDIPLYQGLSLEVKKGQTLAVG 1072
Db 1014 IIEKTPILDSYSTEGIKPTLEGNTFENEVENYPTRPDIPLYQGLSLEVKKGQTLAVG 1073
QY 1073 SSGGKSTVWOLLERFYDPLAGSVLJQGEIKHUNOMLRHIGIYSOEPLIFDCSTAE 1132
Db 1074 SSGGKSTVWOLLERFYDPLAGSVLJQGEIKHUNOMLRHIGIYSOEPLIFDCSTAE 1133
QY 1133 IAYGDSRVVSHETMQAAKEANIHETLPEKYNTRVGDKGTOLSGGQKORAIARAL 1192
Db 1134 IAYGDSRVVSHETMQAAKEANIHETLPEKYNTRVGDKGTOLSGGQKORAIARAL 1193
QY 1193 VROPHILLDEATSALDTESEKVVQEALDKAREGTCIVIAHRLSTIONADLIYVFONGK 1252
Db 1194 VROPHILLDEATSALDTESEKVVQEALDKAREGTCIVIAHRLSTIONADLIYVFONGK 1253
QY 1253 VKEHGTQOOLIAOKGIYFSMVSVQAGAKR 1281
Db 1254 VKEHGTQOOLIAOKGIYFSMVSVQAGAKR 1282

```

Search completed: December 9, 2002, 16:36:37
 Job time : 933 secs

GenCore version 5.1.3
Copyright (c) 1993 - 2002 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: December 9, 2002, 16:20:54 ; Search time 142 Seconds

(without alignments)
146.524 Million cell updates/sec

Title: US-09-672-725C-2

Perfect score: 6477
Sequence: 1 MDPEGRKSGSAEKNFWKMGK.....LTAOKGIFGMSVQAGAKR 1281

Scoring table:

BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 103943 seqs, 16242309 residues

Total number of hits satisfying chosen parameters: 103943

Minimum DB seq length: 0
Maximum DB seq length: 200000000

Post-processing: Minimum Match 08
Maximum Match 100%

Listing first 45 summaries

Database : Published Applications, AA:*

- 1: /cgn2_6/ptodata/2/pubpaa/US08_NEW_PUB.pep.*
- 2: /cgn2_6/ptodata/2/pubpaa/PCT_NEW_PUB.pep.*
- 3: /cgn2_6/ptodata/2/pubpaa/US06_NEW_PUB.pep.*
- 4: /cgn2_6/ptodata/2/pubpaa/US06_PUBCOMB.pep.*
- 5: /cgn2_6/ptodata/2/pubpaa/US07_NEW_PUB.pep.*
- 6: /cgn2_6/ptodata/2/pubpaa/US07_PUBCOMB.pep.*
- 7: /cgn2_6/ptodata/2/pubpaa/PCTUS_PUBCOMB.pep.*
- 8: /cgn2_6/ptodata/2/pubpaa/US08_PUBCOMB.pep.*
- 9: /cgn2_6/ptodata/2/pubpaa/US09_NEW_PUB.pep.*
- 10: /cgn2_6/ptodata/2/pubpaa/US09_PUBCOMB.pep.*
- 11: /cgn2_6/ptodata/2/pubpaa/US10_NEW_PUB.pep.*
- 12: /cgn2_6/ptodata/2/pubpaa/US10_PUBCOMB.pep.*
- 13: /cgn2_6/ptodata/2/pubpaa/US60_NEW_PUB.pep.*
- 14: /cgn2_6/ptodata/2/pubpaa/US60_PUBCOMB.pep.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	6420.5	99.1	1280	9 US-10-044-671-2	Sequence 2, Appli
2	5861	90.5	1280	10 US-10-072-621-7	Sequence 7, Appli
3	5861	90.5	1280	10 US-09-866-866A-2	Sequence 2, Appli
4	5852	90.4	1280	10 US-09-866-866A-4	Sequence 4, Appli
5	5636	87.0	1276	10 US-09-866-866A-8	Sequence 8, Appli
6	5616	86.7	1272	10 US-09-769-097-4	Sequence 4, Appli
7	5606	86.6	1272	10 US-09-866-866A-6	Sequence 6, Appli
8	5296	81.8	1276	10 US-09-866-866A-6	Sequence 2, Appli
9	3549	54.8	1195	10 US-09-873-409-5	Sequence 5, Appli
10	3525.5	54.4	1222	10 US-09-873-409-5	Sequence 6, Appli
11	3112.5	48.1	1058	10 US-09-873-409-4	Sequence 4, Appli
12	2805.5	43.3	1275	10 US-09-873-409-4	Sequence 6, Appli
13	2481	38.3	1334	9 US-09-758-828-2	Sequence 2, Appli
14	2442	37.1	812	10 US-09-873-409-2	Sequence 3, Appli
15	1950	30.1	1394	9 US-10-101-388-3	Sequence 1, Appli
16	1870	28.9	659	10 US-09-873-409-1	Sequence 8, Appli
17	1576	24.3	514	10 US-09-873-409-8	Sequence 7, Appli
18	1552.5	24.0	541	10 US-09-873-409-7	Sequence 14006, A
19	836	12.9	582	10 US-09-815-242-14006	

20	825	12.7	582	10 US-09-815-242-10115	Sequence 10115, A
21	820	12.7	766	9 US-10-072-621-6	Sequence 6, Appli
22	816	12.6	766	10 US-09-795-693-17	Sequence 17, Appli
23	815	12.6	587	10 US-09-815-242-10960	Sequence 10960, A
24	731	11.3	1261	10 US-09-759-143-538	Sequence 538, App
25	731	11.3	1261	10 US-09-780-669-538	Sequence 538, App
26	731	11.3	1261	10 US-09-822-827-537	Sequence 537, App
27	709.5	11.0	603	10 US-09-815-242-5202	Sequence 5202, App
28	695.5	10.7	578	10 US-09-815-242-12501	Sequence 12501, A
29	687	10.6	748	9 US-09-870-759-41	Sequence 41, Appli
30	672	10.4	640	9 US-09-970-059-9	Sequence 9, Appli
31	664	10.3	571	10 US-09-815-242-10619	Sequence 10619, A
32	661	10.2	574	10 US-09-953-688A-1	Sequence 1, Appli
33	661	10.2	836	10 US-09-953-688A-7	Sequence 7, Appli
34	658	10.2	621	10 US-09-953-259-2	Sequence 2, Appli
35	655	10.1	1228	10 US-09-759-143-537	Sequence 537, App
36	655	10.1	1228	10 US-09-780-669-537	Sequence 537, App
37	655	10.1	1228	10 US-09-822-827-537	Sequence 537, App
38	641.5	9.9	583	10 US-09-815-242-13458	Sequence 13458, A
39	636.5	9.8	339	9 US-09-764-884-25	Sequence 25, Appli
40	618	9.5	353	9 US-09-764-884-3	Sequence 3, Appli
41	596	9.2	1480	9 US-09-568-756-2	Sequence 2, Appli
42	544	8.4	551	10 US-09-815-242-11394	Sequence 11394, A
43	528.5	8.2	400	10 US-09-765-272-190	Sequence 190, App
44	523.5	8.1	760	10 US-09-833-017-26	Sequence 26, Appli
45	522.5	8.1	659	10 US-09-841-132-497	Sequence 497, App

ALIGNMENTS

RESULT 1
US-10-044-671-2
Sequence 2, Application US/10044671
Patent No. US2002017147A1
GENERAL INFORMATION:
APPLICANT: Washington State University Research Foundation
APPLICANT: Mealey, Katrina
TITLE OR INVENTION: BENTLEY, Steven
FILE REFERENCE: MDRI VARIANTS AND METHODS FOR THEIR USE
CURRENT APPLICATION NUMBER: US/10/044,671
CURRENT FILING DATE: 2002-01-10
PRIOR APPLICATION NUMBER: US 60/261,578
PRIOR FILING DATE: 2001-01-12
PRIOR APPLICATION NUMBER: US 60/314,829
PRIOR FILING DATE: 2001-08-24
NUMBER OF SEQ ID NOS: 10
SOFTWARE: PatentIn version 3.1
SEQ ID NO 2
LENGTH: 1280
TYPE: PRT
ORGANISM: Canis familiaris
US-10-044-671-2

Query Match 99.1%; Score 6420.5; DB 9; Length 1280;
Best Local Similarity 99.3%; Pred No. 0;
Matches 1272; Conservative 4; Mismatches 4; Indels 1; Gaps 1;

QY 1 MDPEGRKSGSAEKNFWKMGKSKKKEKKPTVSTFAMFRYSNLDRLVMTGMAII 60
|||||
DB 1 MDPEGRKSGSAEKNFWKMGKSKK-EKREKKPTVSTFAMFRYSNLDRLVMTGMAII 59
|||||

QY 61 HGAALPLMLLVGNMDDPANNNGISRNKTFEPVINESTNTQHFHNLEEMTVYVY 120
|||||
DB 60 HGAALPLMLLVGNMDDPANNNGISRNKTFEPVINESTNTQHFHNLEEMTVYVY 119
|||||

QY 121 SGIGAGVLAAYIYVSWCLAGROILKIRKOFHAIKROEIGFVDVDELWTRITDD 180
|||||
DB 120 SGIGAGVLAAYIYVSWCLAGROILKIRKOFHAIKROEIGFVDVDELWTRITDD 179
|||||

QY 181 VKINGIGDKIGMFHSTATFFGFIYGFIRGKMLTVLVAISPVGLSAAIWAII 240
|||||


```

Db 658 SSSLIKRSTRRSVRSQAODRKLSTKEALDESIPPVSFWRIMKLNLTMPYFVVGVGC 717
Qy 720 AINGGLOPAFISIFSRIGIFTRDEDEPETKRONSNNEVLFVLGISTITFFLOGFTF 779
Db 718 AINGGLOPAFALIFSKITIGVTRIDDEPETKRONSNLFSLFLALGISTITFFLOGFTF 777
Qy 780 GKAGEILLKRLRYVFRSMRLRODVSWFDDPKNTGALTTRLANDAAOVGAIGSRLAVIT 839
Db 778 GKAGEILLKRLRYVFRSMRLRODVSWFDDPKNTGALTTRLANDAAOVGAIGSRLAVIT 837
Qy 840 QNINLGTGIIISLYGMQTLTLLLAIVPIIAAGVEMKMSGQALKDKKELEGAGKTA 899
Db 838 QNINLGTGIIISLYGMQTLTLLLAIVPIIAAGVEMKMSGQALKDKKELEGAGKTA 897
Qy 900 TEALENFRVVSILTRQCKFEYMTAOSLOVYRNSLKAHIFGVSPSTIQAMMYFYACGF 959
Db 898 TEALENFRVVSILTRQCKFEYMTAOSLOVYRNSLKAHIFGVSPSTIQAMMYFYACGF 957
Qy 960 RFGAYLVANEFNFOVLLVFSALVFGAMAVGVSSFPADYAKAKVSAHVIMIIEKSP 1019
Db 958 RFGAYLVANEFNFOVLLVFSALVFGAMAVGVSSFPADYAKAKVSAHVIMIIEKSP 1017
Qy 1020 IDSYPHGLKPNLTLEGNVTENEVFNYPTRPDIPVLOGISLEVKKGQTLALVSGSGCKS 1079
Db 1018 IDSYPHGLKPNLTLEGNVTENEVFNYPTRPDIPVLOGISLEVKKGQTLALVSGSGCKS 1077
Qy 1080 TVVOLLERYPDLPLAGSVLLDGKEIKRLNVOMLRAHLGIVSOEPIIFDCSIENIAYGDN 1139
Db 1078 TVVOLLERYPDLPLAGSVLLDGKEIKRLNVOMLRAHLGIVSOEPIIFDCSIENIAYGDN 1137
Qy 1140 RVVSHIEIMQAKEANIHFIETLPEKYNTRVGDGQTLSSGOKORAIARALVRPHIL 1199
Db 1138 RVVSHIEIMQAKEANIHFIETLPEKYNTRVGDGQTLSSGOKORAIARALVRPHIL 1197
Qy 1200 LIDEATSALDTESEKVVODALDKAREGRTCIYIAHRLSTIONMADLIYVFOGKXKEHGT 1259
Db 1198 LIDEATSALDTESEKVVODALDKAREGRTCIYIAHRLSTIONMADLIYVFOGKXKEHGT 1257
Qy 1260 QOLLAOKGIYFSMVSVOAGAKR 1281
Db 1258 QOLLAOKGIYFSMVSVOAGAKR 1279

```

RESULT 3

```

; US-09-866-866A-2
; Sequence 2, Application US/09866866A
; Patent No. US20020102244A1
; GENERAL INFORMATION:
; APPLICANT: Sorrentino, Brian
; TITLE OF INVENTION: A Method of Identifying and/or Isolating Stem Cells
; FILE REFERENCE: 1340-1-021CIP2
; CURRENT APPLICATION NUMBER: US/09/866, 866A
; PRIOR APPLICATION NUMBER: 2001-08-30
; PRIOR FILING DATE: 2000-05-31
; PRIOR APPLICATION NUMBER: PCT/US99/11825
; PRIOR FILING DATE: 1999-05-27
; PRIOR APPLICATION NUMBER: 60/086, 988
; PRIOR FILING DATE: 1998-05-28
; NUMBER OF SEQ ID NOS: 27
; SOFTWARE: PatentIn version 3.0
; SEQ ID NO 2
; LENGTH: 1280
; TYPE: PRT
; ORGANISM: Homo sapiens
; US-09-866-866A-2

```

```

Query Match 90.5%; Score 5861; DB 10; Length 1280;
Best Local Similarity 90.7%; Pred. No. 0;
Matches 1163; Conservative 55; Mismatches 60; Indels 4; Gaps 4;
Qy 1 MDEGGRRGSA-EKNFMKGKSKKNEKKEKPYSTFAMFYSNMDLRLTYLVGTMAAI 59

```

```

Db 1 MDLGGDRNGAKAKKNEFKLNKSEK-DKKEKCTVYSFMSFRSNMDLRLTYLVGTMAAI 59
Qy 60 IHGAALPLMIVENMMDSPANAGISNKTFFPIYINESTNNTOHPINHLEEMTYAY 119
Db 60 IHGAGLPLMLVGEEMTDIFANAG-NLEDLMSNITKRSIDINDGPFMN-LEEDMTRYAY 117
Qy 120 YSGIAGVLVAAVYIOVSFWCLAGROTLKTRKOFPAIRKROEIGMDVDVDELNTRLTD 179
Db 118 YSGIAGVLVAAVYIOVSFWCLAGROTLKTRKOFPAIRKROEIGMDVDVDELNTRLTD 177
Qy 180 DVSKINCGIDKIGMFPHSIAFTFTGIVFTGRMKTLLVILAIISPVLSIAIYAKILS 239
Db 178 DVSKINCGIDKIGMFPHSIAFTFTGIVFTGRMKTLLVILAIISPVLSIAIYAKILS 237
Qy 240 SFTDKELLAAYAKAGAAVEYLAIRTVIAFGQKKELEKRYNNLEBANGIGIKRATANI 299
Db 238 SFTDKELLAAYAKAGAAVEYLAIRTVIAFGQKKELEKRYNNLEBANGIGIKRATANI 297
Qy 300 SIGAFLLLIYASYALAFWYGTSLVLSSEYSIGQVLTFFSVYLGAFSIGOASPIEAFAN 359
Db 298 SIGAFLLLIYASYALAFWYGTSLVLSSEYSIGQVLTFFSVYLGAFSIGOASPIEAFAN 357
Qy 360 ARGAAYEIKIIDNKPSIDISYKSGHKPDNITGNLEFNHVSYPKREKVLKGLNLKV 419
Db 358 ARGAAYEIKIIDNKPSIDISYKSGHKPDNITGNLEFNHVSYPKREKVLKGLNLKV 417
Qy 420 QSGQTVLVGNSGCGKSTTVOLMQRLYDPTDGMVINDQDRTINVRHLREITGVSOEP 479
Db 418 QSGQTVLVGNSGCGKSTTVOLMQRLYDPTDGMVINDQDRTINVRHLREITGVSOEP 477
Qy 480 VLEATTIANIYKGENYTMDEIEKAYVEANAYDFIMKLPKFDVLVGERCAQLSGGQK 539
Db 478 VLEATTIANIYKGENYTMDEIEKAYVEANAYDFIMKLPKFDVLVGERCAQLSGGQK 537
Qy 540 RLAIARALVRNPKIILLDEATSALDTESEAVVOALDKARGRTTIYIAHRLSVRADY 599
Db 538 RLAIARALVRNPKIILLDEATSALDTESEAVVOALDKARGRTTIYIAHRLSVRADY 597
Qy 600 IAGFDGYIVKGNHDELMKEKGIYFKLVMTQRTGNEIELENATGESKSDALEMSPKD 659
Db 598 IAGFDGYIVKGNHDELMKEKGIYFKLVMTQRTGNEIELENATGESKSDALEMSPKD 657
Qy 660 SSSLIKRSTRRSIIHAQODRKLGTREDLNEVFPVSWRILKLNSTWPFVVGVC 719
Db 658 SSSLIKRSTRRSVRSQAODRKLSTKEALDESIPPVSFWRIMKLNLTMPYFVVGVGC 717
Qy 720 AINGGLOPAFISIFSRIGIFTRDEDEPETKRONSNNEVLFVLGISTITFFLOGFTF 779
Db 718 AINGGLOPAFALIFSKITIGVTRIDDEPETKRONSNLFSLFLALGISTITFFLOGFTF 777
Qy 780 GKAGEILLKRLRYVFRSMRLRODVSWFDDPKNTGALTTRLANDAAOVGAIGSRLAVIT 839
Db 778 GKAGEILLKRLRYVFRSMRLRODVSWFDDPKNTGALTTRLANDAAOVGAIGSRLAVIT 837
Qy 840 QNINLGTGIIISLYGMQTLTLLLAIVPIIAAGVEMKMSGQALKDKKELEGAGKTA 899
Db 838 QNINLGTGIIISLYGMQTLTLLLAIVPIIAAGVEMKMSGQALKDKKELEGAGKTA 897
Qy 900 TEALENFRVVSILTRQCKFEYMTAOSLOVYRNSLKAHIFGVSPSTIQAMMYFYACGF 959
Db 898 TEALENFRVVSILTRQCKFEYMTAOSLOVYRNSLKAHIFGVSPSTIQAMMYFYACGF 957
Qy 960 RFGAYLVANEFNFOVLLVFSALVFGAMAVGVSSFPADYAKAKVSAHVIMIIEKSP 1019
Db 958 RFGAYLVANEFNFOVLLVFSALVFGAMAVGVSSFPADYAKAKVSAHVIMIIEKSP 1017
Qy 1020 IDSYPHGLKPNLTLEGNVTENEVFNYPTRPDIPVLOGISLEVKKGQTLALVSGSGCKS 1079
Db 1018 IDSYPHGLKPNLTLEGNVTENEVFNYPTRPDIPVLOGISLEVKKGQTLALVSGSGCKS 1077
Qy 1080 TVVOLLERYPDLPLAGSVLLDGKEIKRLNVOMLRAHLGIVSOEPIIFDCSIENIAYGDN 1139

```

```
Db 1078 TVVOLLERFYDPLAGVLLDGEIKRLNVOMLRAHGIYSOEPIFDCSIENIAVGDNS 1137
Qy 1140 RVVSHHEIMOAKKANIHIEITLPEKYNTRVGDKGTOLSGGOKORIAIARALVROPAIL 1199
Db 1138 RVVSOEIVRAKAKENAHIEITLPEKYNTRVGDKGTOLSGGOKORIAIARALVROPAIL 1197
Qy 1200 LLDENATSDTESEKVVQVQALDKARREGRTCIYIAHRLSTIONADLIYVQNGVKEHGT 1259
Db 1198 LLDENATSDTESEKVVQVQALDKARREGRTCIYIAHRLSTIONADLIYVQNGVKEHGT 1257
Qy 1260 QOLLAQGIYFSMYSVOAGAKR 1281
Db 1258 QOLLAQGIYFSMYSVOAGAKR 1279

RESULT 4
US-09-866-4
; Sequence 4, Application US/09866866A
; Patent No. US2002010224A1
; GENERAL INFORMATION:
; APPLICANT: Sorrentino, Brian
; APPLICANT: Schuetz, John
; TITLE OF INVENTION: A Method of Identifying and/or Isolating Stem Cells
; FILE REFERENCE: 1340-1-02ICP2
; CURRENT APPLICATION NUMBER: US/09/866, 866A
; CURRENT FILING DATE: 2001-08-30
; PRIOR APPLICATION NUMBER: 09/584, 586
; PRIOR FILING DATE: 2000-05-31
; PRIOR APPLICATION NUMBER: PCT/US99/11825
; PRIOR FILING DATE: 1999-05-27
; PRIOR APPLICATION NUMBER: 60/086, 988
; PRIOR FILING DATE: 1998-05-28
; NUMBER OF SEQ ID NOS: 27
; SOFTWARE: PatentIn version 3.0
; SEQ ID NO 4
; LENGTH: 1280
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-866-866A-4

Query Match 90.4%; Score 5852; DB 10; Length 1280;
Best Local Similarity 90.6%; Pred. No. 0;
Matches 1162; Conservative 55; Mismatches 61; Indels 4; Gaps 4;

Qy 1 MDPEGGRKGS- EKNFWKMGKSKKNEKKKFTVSTFAMFRYSNMDDLVMNGMTMAI 59
Db 1 MDEGDRNGGAKKKKNNKSEK-DKKEKKPTVSFMSFRYSNWLDKLVVNGTLAAI 59
Qy 60 IHGAALPLMLVFGNMTDFRANGISNKTPEVITNESITNNTOHFINLHEEMTYAY 119
Db 60 IHGAGLPMLVFGEMTDIFRANG-NLEDLSMTNRSOLDNIGFPMN-LEEDMTRAYAY 117
Qy 120 YSGIGAGVLAAYIQVSWFMCIAAGROILKIRKOFYAHIMROEIGMFDVHVGELNRLD 179
Db 118 YSGIGAGVLAAYIQVSWFMCIAAGROILKIRKOFYAHIMROEIGMFDVHVGELNRLD 177
Qy 180 DVSRTINEGIGDKIGMFRHSATFTFTGTVGFTTRGMKLTIVLAIISPLGLSAIIMAKIS 239
Db 178 DVSRTINEGIGDKIGMFRHSATFTFTGTVGFTTRGMKLTIVLAIISPLGLSAIIMAKIS 237
Qy 240 SPTDEKLAAYKAGAVAEVLAIRTYIAFGQKKELEERYNKNLEAKGIGIKKAITANI 299
Db 238 SPTDEKLAAYKAGAVAEVLAIRTYIAFGQKKELEERYNKNLEAKGIGIKKAITANI 297
Qy 300 SIGAFLIYASALAFMYGTSLVLSSEYSIGQVLYVFSVLGAFSIGASPIEAFAN 359
Db 298 SIGAFLIYASALAFMYGTSLVLSSEYSIGQVLYVFSVLGAFSIGASPIEAFAN 357
Qy 360 AGAAVEIFKTIIDNKPSIDSYSGSKHDPDNKGNLEFNHVSYPSSREKVKILGLMLKY 419
Db 358 AGAAVEIFKTIIDNKPSIDSYSGSKHDPDNKGNLEFNHVSYPSSREKVKILGLMLKY 417
Qy 420 OSGQTVALVNSGCGKSTTVOLMORLYDPDGMVCIDGQDIRTINVRKHLREITGVSOEP 479
```

```
Db 418 OSGQTVALVNSGCGKSTTVOLMORLYDPDGMVSVQDQDIRTINVRKHLREITGVSOEP 477
Qy 480 VLFATTAENIRYGRVNTWDEIEKAVKEANAYDFIMKLPNKFDTLVGSGAQLSGGOK 539
Db 478 VLFATTAENIRYGRVNTWDEIEKAVKEANAYDFIMKLPNKFDTLVGSGAQLSGGOK 537
Qy 540 RIATARALVNPRTILLDEATSDTESEKVVQVQALDKARREGRTCIYIAHRLSTIONAD 599
Db 538 RIATARALVNPRTILLDEATSDTESEKVVQVQALDKARREGRTCIYIAHRLSTIONAD 597
Qy 600 IAGPDGVIVKGNHDELMKKEGIFYFLVMTQTRGNEIELENATGESKESDALEMSND 659
Db 598 IAGPDGVIVKGNHDELMKKEGIFYFLVMTQTRGNEIELENATGESKESDALEMSND 657
Qy 660 SSSSLIKRRSTRSRIHAPQDGRKLTREDNEVNPVPSWRILKLNSTPEYFVGVIC 719
Db 658 SSSSLIKRRSTRSRIHAPQDGRKLTREDNEVNPVPSWRILKLNSTPEYFVGVIC 717
Qy 720 AIIINGLOPAPSIIFSRIGIFTRDEDEPERKONSNNFSLFLALGIIISFTFELQGFTE 779
Db 718 AIIINGLOPAPSIIFSRIGIFTRDEDEPERKONSNNFSLFLALGIIISFTFELQGFTE 777
Qy 780 GRAGEITKRLRYVFRSMLRDVSWEDDPKNTTGALTTRLANDAQAQVAGISRLAVIT 839
Db 778 GRAGEITKRLRYVFRSMLRDVSWEDDPKNTTGALTTRLANDAQAQVAGISRLAVIT 837
Qy 840 ONIANIGTIIISLYGQWLTLLIATVPIAIAIGVEMKLSQALDKKELEGACKIA 899
Db 838 ONIANIGTIIISLYGQWLTLLIATVPIAIAIGVEMKLSQALDKKELEGACKIA 897
Qy 900 TEAIENTFRVSLTRKREKRYMAOSLOVPRNSIRKAHIFGVSEFSTQAMMYSVAGCF 959
Db 898 TEAIENTFRVSLTRKREKRYMAOSLOVPRNSIRKAHIFGVSEFSTQAMMYSVAGCF 957
Qy 960 REGAYIVANEENFQDVLVFSAYIFGMAVGOVSSPADVAKKISAHIMIEIKTPL 1019
Db 958 REGAYIVANEENFQDVLVFSAYIFGMAVGOVSSPADVAKKISAHIMIEIKTPL 1017
Qy 1020 IDSYPHGLKPNLTLEGNVTFENEVNYPTRPDIPLYGLGISEVKKGOTLALVSSGCKS 1079
Db 1018 IDSYPHGLKPNLTLEGNVTFENEVNYPTRPDIPLYGLGISEVKKGOTLALVSSGCKS 1077
Qy 1080 TVVOLLERFYDPLAGVLLDGEIKRLNVOMLRAHGIYSOEPIFDCSIENIAVGDNS 1139
Db 1078 TVVOLLERFYDPLAGVLLDGEIKRLNVOMLRAHGIYSOEPIFDCSIENIAVGDNS 1137
Qy 1140 RVVSHHEIMOAKKANIHIEITLPEKYNTRVGDKGTOLSGGOKORIAIARALVROPAIL 1199
Db 1138 RVVSOEIVRAKAKENAHIEITLPEKYNTRVGDKGTOLSGGOKORIAIARALVROPAIL 1197
Qy 1200 LLDENATSDTESEKVVQVQALDKARREGRTCIYIAHRLSTIONADLIYVQNGVKEHGT 1259
Db 1198 LLDENATSDTESEKVVQVQALDKARREGRTCIYIAHRLSTIONADLIYVQNGVKEHGT 1257
Qy 1260 QOLLAQGIYFSMYSVOAGAKR 1281
Db 1258 QOLLAQGIYFSMYSVOAGAKR 1279

RESULT 5
US-09-866-8
; Sequence 8, Application US/09866866A
; Patent No. US2002010224A1
; GENERAL INFORMATION:
; APPLICANT: Sorrentino, Brian
; APPLICANT: Schuetz, John
; TITLE OF INVENTION: A Method of Identifying and/or Isolating Stem Cells
; FILE REFERENCE: 1340-1-02ICP2
; CURRENT APPLICATION NUMBER: US/09/866, 866A
; CURRENT FILING DATE: 2001-08-30
; PRIOR APPLICATION NUMBER: 09/584, 586
; PRIOR FILING DATE: 2000-05-31
```

; PRIOR APPLICATION NUMBER: PCT/US99/11825
 ; PRIOR FILING DATE: 1999-05-27
 ; PRIOR APPLICATION NUMBER: 60/086, 988
 ; PRIOR FILING DATE: 1998-05-28
 ; NUMBER OF SEQ ID NOS: 27
 ; SOFTWARE: PatentIn version 3.0
 ; SEQ ID NO 8
 ; LENGTH: 1276
 ; TYPE: PRF
 ; ORGANISM: Mus musculus
 us-09-866-866A-8

Query Match 87.0%; Score 5636; DB 10; Length 1276;
 Best Local Similarity 87.0%; Pred. No. 0;
 Matches 115; Conservative 75; Mismatches 84; Indels 8; Gaps 4;

QY 1 MDEGGRKSAEKNFMGKSKKKKKRPTVTFAMFRKSNMLDRLYMLVGTMAII 60
 DB 1 MELEEDLKGADKNFSKMGKSKK-EKKKKPAVSVLTFMFRAGWLDRLYMLVGTMAII 59
 QY 61 HGAALPLMLVFGNMTDSFANNG-ISRNKTPVLIINESITNTQHFINHLEEMTYAYY 119
 DB 60 HGAALPLMLIFGDMTDSFASVGNVSKNST-----NMSADKRAHFAR-LDEEMTYAYY 113
 QY 120 YSGIGAVLYAAYIOVSEFCLAGROILIKRQFTHAMROEIGFVDVHVGELNRLTD 179
 DB 114 YTGIGAVLYAAYIOVSEFCLAGROILIKRQFTHAMROEIGFVDVHVGELNRLTD 173
 QY 180 DVSKINEGIDRIKGFHSHIATFTGFTVGTGKTLVLTLAISPVYIGISAATWAKILS 239
 DB 174 DVSKINEGIDRIKGFHSHIATFTGFTVGTGKTLVLTLAISPVYIGISAATWAKILS 233
 QY 240 SPTDEKLLAYAKAGAAVEVLAIRTVAEGGKKELEERYKNLDEAKGIGIKKAITANI 239
 DB 234 SPTDEKLLAYAKAGAAVEVLAIRTVAEGGKKELEERYKNLDEAKGIGIKKAITANI 233
 QY 300 SIGAFLILYASYALAFYGTSLVSEYSIGOVLTVPFESVLIGAFSIGASSTIEFAN 359
 DB 294 SMGAFLILYASYALAFYGTSLVSEYSIGOVLTVPFESVLIGAFSIGASSTIEFAN 353
 QY 360 ARGAAVEYLFKTIIDNKPISIDTSKSGHKPDNIKNGLEFRNHFVSPSRKVKIILKGLTKY 419
 DB 354 ARGAAVEYLFKTIIDNKPISIDTSKSGHKPDNIKNGLEFRNHFVSPSRKVKIILKGLTKY 413
 QY 420 QSGQVALVNGSGCGKSTTYVQIMORLDPDGMVCTIDGQIRITINVRHLEITGVVSOEP 479
 DB 414 KSGQVALVNGSGCGKSTTYVQIMORLDPDGMVCTIDGQIRITINVRHLEITGVVSOEP 473
 QY 480 VLFATTIAENIRYGRENVYMDIEKAVKANAYDFIMKLPNKFDTLVGERGAOLSGGOKO 539
 DB 474 VLFATTIAENIRYGRENVYMDIEKAVKANAYDFIMKLPNKFDTLVGERGAOLSGGOKO 533
 QY 540 RIATAARLVPRPKILLDEATSAIDPESEAVNOVALDKARKGTTIVIAHRLSTVKNADY 599
 DB 534 RIATAARLVPRPKILLDEATSAIDPESEAVNOVALDKARKGTTIVIAHRLSTVKNADY 593
 QY 600 IAGDDGVIVEKGNHDEIMKEKGIYFKLYVTQTRGNEIELEJNATGSKESALTEMSPD 659
 DB 594 IAGDDGVIVEKGNHDEIMKEKGIYFKLYVTQTRGNEIELEJNATGSKESALTEMSPD 653
 QY 660 SSSLLIKRSTRSIAHPOGODRKLTGEDLNENPVPSFWIRILKINSTEMPPYVVGIFC 719
 DB 654 SSSLLIKRSTRSIAHPOGODRKLTGEDLNENPVPSFWIRILKINSTEMPPYVVGIFC 713
 QY 720 AIIINGLOPAFSIIFSRNIGIFRDEDPETKRONSMFVLVLGIVTSITFTFLOGTTF 779
 DB 714 AIIINGLOPAFSIIFSRNIGIFRDEDPETKRONSMFVLVLGIVTSITFTFLOGTTF 773
 QY 780 GRKGEILLTRKLRMYRSMLRQDVSWFDDPKMTGTALTRLANDAAQVKGAIISRLAIVT 839
 DB 774 GRKGEILLTRKLRMYRSMLRQDVSWFDDPKMTGTALTRLANDAAQVKGAIISRLAIVT 833
 QY 840 QNTANLGTGIIISLYGWQLTLLLAIVPIIAIAGVEMKMLSGOALKDKKLEBAGAKIA 899

DB 834 QNINLGTGIIISLYGWQLTLLLAIVPIIAIAGVEMKMLSGOALKDKKLEBAGAKIA 893
 QY 900 TEALENFRYVSLTREKFEYMTAOSLQVRYNSLRKAHIFGVSESTQAMMYFSYAGCF 959
 DB 894 TEALENFRYVSLTREKFEYMTAOSLQVRYNSLRKAHIFGVSESTQAMMYFSYAGCF 953
 QY 960 RFGALYVANEFMNFODILVFSAIVFGAMAVGVSSFADPAKATVSAASHIIRIIEKTP 1019
 DB 954 RFGALYVANEFMNFODILVFSAIVFGAMAVGVSSFADPAKATVSAASHIIRIIEKTP 1013
 QY 1020 IDSYSPGKLPNTLEGNVTNEVFNYPTRPDIPVLOGLSLEKKGQTLATVSSGCGKS 1079
 DB 1014 IDSYSPGKLPNTLEGNVTNEVFNYPTRPDIPVLOGLSLEKKGQTLATVSSGCGKS 1073
 QY 1080 TVVQLERFDPPLAGSVLIDGKEIKHLNQMRAHLGIVSOPIIFDCSIAENIAYGNS 1139
 DB 1074 TVVQLERFDPPLAGSVLIDGKEIKHLNQMRAHLGIVSOPIIFDCSIAENIAYGNS 1133
 QY 1140 RVVSHHEIQAQAEANIHFIETLEPKYNTRVGDKGTOLSGGOKORIAIARALYRPHIL 1199
 DB 1134 RVVSHHEIQAQAEANIHFIETLEPKYNTRVGDKGTOLSGGOKORIAIARALYRPHIL 1193
 QY 1200 LIDEATSAIDTESKYNQVADLKAERGRCIYIAHRLSTIONADLIVFGONGKVEGTH 1259
 DB 1194 LIDEATSAIDTESKYNQVADLKAERGRCIYIAHRLSTIONADLIVFGONGKVEGTH 1253
 QY 1260 QOLLAOKGYFSMSVQAGAKR 1281
 DB 1254 QOLLAOKGYFSMSVQAGAKR 1275

RESULT 6

us-09-769-097-4
 ; Sequence 4, Application US/09769097
 ; Patent No. US20020055128A1
 ; GENERAL INFORMATION:
 ; APPLICANT: Kimberly Anne Brun
 ; APPLICANT: Richard James Chenery
 ; APPLICANT: Harma Ellens
 ; APPLICANT: John Anthony Feld
 ; APPLICANT: Lin Yue
 ; TITLE OF INVENTION: POLYNUCLEOTIDE AND POLYPEPTIDE SEQUENCES
 ; FILE REFERENCE: GP-50009-C2
 ; CURRENT APPLICATION NUMBER: US/09/769,097
 ; PRIORITY FILING DATE: 2001-01-25
 ; PRIOR APPLICATION NUMBER: 09/208,809
 ; PRIOR FILING DATE: 1998-12-09
 ; PRIOR APPLICATION NUMBER: 09/156,800
 ; PRIOR FILING DATE: 1999-09-17
 ; PRIOR APPLICATION NUMBER: US99/20770
 ; NUMBER OF SEQ ID NOS: 4
 ; SOFTWARE: FastSeq for Windows Version 3.0
 ; SEQ ID NO 4
 ; LENGTH: 1272
 ; TYPE: PRF
 ; ORGANISM: RATTUS RATTUS
 us-09-769-097-4

Query Match 86.7%; Score 5616; DB 10; Length 1272;
 Best Local Similarity 86.3%; Pred. No. 0;
 Matches 1105; Conservative 81; Mismatches 85; Indels 10; Gaps 2;

QY 1 MDEGGRKSAEKNFMGKSKKKKKRPTVTFAMFRKSNMLDRLYMLVGTMAII 60
 DB 1 MELEEDLKGADKNFSKMGKSKK-EKKKKPAVSVLTFMFRAGWLDRLYMLVGTMAII 59
 QY 61 HGAALPLMLVFGNMTDSFANNGISRNTFPVLIINESITNTQHFINHLEEMTYAYY 120
 DB 60 HGAALPLMLVFGNMTDSFANNGISNRNTFPVLIINESITNTQHFINHLEEMTYAYY 110


```

/
/   PRIOR FILING DATE: 1998-05-28
/   NUMBER OF SEQ ID NOS: 27
/   SOFTWARE: PatentIn version 3.0
/   SEQ ID NO 6
/   LENGTH: 1276
/   TYPE: PRT
/   ORGANISM: Mus musculus
US-09-866-866A-6

```

Query Match	Best local Similarity	81.8%	Score 5296;	DB 10;	Length 1276;
Matches 1029;	Conservative 130;	Mismatches 115;	Indels 8;	Gaps 5	
QY	1	MDEGGRKSAKNNWKKGKSKKNEKKKPEYSTAMFRYSNWLDRLMYLGYMAAII	60		
Db	1	MEFEENLGRADKNKSKMGKSKK- EKREKKPAGVFGRFYADWDKLCMILGTAAYI	59		
QY	61	HCALPLMLVFGNNTDSFANAGISRKKTFPVIIINESITNNTOHFII- HLEEMTYAYI	119		
Db	60	HETLLPLMLVFGNNTDSFETKAAS- --IIPSTINSGPSTLLISNSSLEEMLAYAYI	116		
QY	120	YSGIAGVLVAAYIOVSPWCLAAAROLIKRKQEPHAIHQEIGRFDVHDVGEIINTRLD	179		
Db	117	YTGIGAGVLVAAYIOVSLMCLAAGRQHLKTRQKPFHAIHQEIGRFDVHDVGEIINTRLD	176		
QY	160	DYSKINESTGKICMFPFHSIATPEFTGPIVGFSTGKWLITVLIASPVLELSAIAKILIS	239		
Db	177	DYSKINDIGIKIMFPOSITTFELAGIIGFISGWLITVLIILAVSPILISALMAKAVLT	236		
QY	240	SFTPKELIAVAKAVAEVLAIRVIAAFGGCKKLEKRYNNKLEBAKGIGIKKAITANI	299		
Db	237	SFTPKELIAVAKAVAEVLAIRVIAAFGGCKKLEKRYNNKLEBAKNVGIKKAITYASI	296		
QY	300	SIGAAFLIIVASVALAEWGTSLVLSSESYIGCVLIVFVFSGVIGAFISQASPSIEAN	359		
Db	297	SIGAIYVLVYASVALAEWGTSLVLSSESYIGCVLIVFVFSGVIGAFISQASPSIEAN	356		
QY	360	ARGAAYIEFIIDNKKPSISYSKSGKRPNIKGNLEFKNVEHSYSRREKVTILGNLKV	419		
Db	357	ARGAFEFITIIDNESIDFSFTKGKRPDSIKGNLEFKNVEHSYSRREKVTILGNLKV	416		
QY	420	QSGQFVALVGNSSCGKSTTVOLMQRIYDPTDQVCIIDQDRIITINVRILREITVSGQEP	479		
Db	417	KSGQFVALVGNSSCGKSTTVOLMQRIYDPTDQVCIIDQDRIITINVRILREITVSGQEP	476		
QY	480	VLEPATTIAENIRGRNNTVDELIEKAVKEANAADFIMKLPRKFDVLVGERGAQLSGQOK	539		
Db	477	VLEPATTIAENIRGRNNTVDELIEKAVKEANAADFIMKLPRKFDVLVGERGAQLSGQOK	536		
QY	540	RIAAARALVRNKIILLDEATSAIDTSEEAVOVALAKARKGRITTYIAHRLSTVRADY	599		
Db	537	RIAAARALVRNKIILLDEATSAIDTSEEAVOVALAKARKGRITTYIAHRLSTVRADY	596		
QY	600	INGFDGVIYVEGNHDELMKEGIIYFKLVIMQTRGNTELEMAATGESKESDALSMPKD	659		
Db	597	INGFDGVIYVEGNHDELMKEGIIYFKLVIMQTRGNTELEMAATGESKESDALSMPKD	656		
QY	660	SGSSILIKRRSTSRSHAPQODRKLGIRKEDINENVPVSWRILKILANSTEMPEYVVGIFC	719		
Db	657	SGSSILIKRRSTSRSHAPQODRKLGIRKEDINENVPVSWRILKILANSTEMPEYVVGIFC	715		
QY	720	AIINGDLOPASIIFSRITIGITFTRDEDEPTRLKONSNMFVSLFVLGIIISITFIFLOGFTF	779		
Db	716	AIINCIQLOPVAIVFSRIYGVFSRDDDEHETKRONCNLFSLFVLVGLISFYIYFPOGFTF	775		
QY	780	KGAGEIILKRLRYVWFWFSMLRODYSWDDDKKNTGALTTRLANDAAQVGAIGSRVAVIT	839		
Db	776	KGAGEIILKRLRYVWFWFSMLRODYSWDDDKKNTGALTTRLANDAAQVGAIGSRVAVIT	835		
QY	840	ONIANLGTGIIISLYGMQTLTLLLAIVPILIAAGVVEKMKSGQALDKRKELEBAGTIA	899		
Db	836	ONIANLGTGIIISLYGMQTLTLLLAIVPILIAAGVVEKMKSGQALDKRKELEBAGTIA	896		

QY	900	TEALENRTVVALSTREBROKFEYNTAOSLOVPYRNSLRKAHIFGVSFTTQAMMTFVAGE	959
Db	896	TEALENRTIVSLTRBQKFEYNTAOSLOVPYRNAMKRAHVGIFSETQAMMTFVAGE	955
QY	960	REGATVLANEFANFOVLLVFSATVFGMAVGOVSSFPADKAKKVAHAIVIMIEKSPD	1019
Db	956	RFGATVLAQOIMTFEENMVLVFSATVFGMAAGNTSSFPADKAKKVASHTIRIEKTPR	1015
QY	1020	IDSTSPHGLPNTLEGNVTNFENEVNYPTRPDI PVYLGSLSEVKKGGTALVYSSGGGKS	1079
Db	1016	IDSTSTEGIKPRTLEGNVKNVGOVNYPTRPDI PVYLGSLSEVKKGGTALVYSSGGGKS	1075
QY	1080	TVVOLLERFPDPLAGSVLLIDGKEIRKLANOMLRHAHIGVISOEPILEFDSIAENIATYDGS	1139
Db	1076	TVVOLLERFPDPMAGSVLEFGKEIKOLANOMLRHAHIGVISOEPILEFDSIAENIATYDGS	1135
QY	1140	RVSHEETMOAKAEANVHHEITFLPERKYNTRVGDKTOLSGGQOKORIAIAALVROPHIL	1199
Db	1136	RAVSHETVIRAKAEANVHGFIDSLPRKYNTRVGDKTOLSGGQOKORIAIAALVROPHIL	1195
QY	1200	ILDEATSAIDTESEKVVOEALDKARERBCITVIAHLSLTIONADLVIFONGKVEHGHT	1259
Db	1196	ILDEATSAIDTESEKVVOEALDKARERBCITVIAHLSLTIONADLVILENGKVEHGHT	1255
QY	1260	QOLLAOKGIYSMVSQVQAKR 1281	
Db	1256	QOLLAOKGIYSM--YQVQAKR 1275	

```

RESULT 9
US-09-873-409-6
: Sequence 6, Application US/09873409
: Patent No. US20020037522A1
GENERAL INFORMATION:
APPLICANT: Frank, Markus
APPLICANT: Sayegh, Mohamed
TITLE OF INVENTION: A Gene Encoding a Multidrug Resistance Human P-glycoprotein
TITLE OF INVENTION: Homologue on Chromosome 7p15-21 and Uses Thereof
FILE REFERENCE: 81994/268611
CURRENT APPLICATION NUMBER: US/09/873,409
CURRENT FILING DATE: 2001-06-05
NUMBER OF SEQ ID NOS: 19
SOFTWARE: PatentIn version 3.0
SEQ ID NO 6
LENGTH: 1195
TYPE: PRT
ORGANISM: Homo sapiens
US-09-873-409-6

```

[illegible][illegible]

```

; TITLE OF INVENTION: Homologue on Chromosome 7p15-21 and Uses thereof
; FILE REFERENCE: 81994/268611
; CURRENT APPLICATION NUMBER: US/09/873.409
; CURRENT FILING DATE: 2001-06-05
; NUMBER OF SEQ ID NOS: 19
; SOFTWARE: PatentIn version 3.0
; SEQ ID NO 5
; LENGTH: 1222
; TYPE: PRT
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: No. US20020037522A1e
; LOCATION: (230)..(230)
; OTHER INFORMATION: Xaa at position 230 represents any L amino acid
US-09-873-409-5

```

```

Query Match          54.4%; Score 3525.5; DB 10; Length 1222;
Best Local Similarity 54.1%; Pred. No. 2.2e-262;
Matches 681; Conservative 238; Mismatches 270; Indels 69; Gaps 9;

```

```

QY 51 MLVGTMAIIGGALPLMLVFNMTDSFANAGISRKTPVILINSITNTQHFINHLE 110
DB 1 MIIIGTASLVNAGACPLMPPLVIGMSDNLSGCL-----VQTNISFRR--- 44
QY 111 EEMTYATYYSGIGAGVVAAYIOVSFWCLAAGROILKIRKOFHAIMROELGMPDVHY 170
DB 45 -----LTLYYGIVGVAALIFGYIOISLIMITTAARQETRIKQFHSVLADQIGMFDSCI 99
QY 171 GELMTRLDDVSKINEGIDKIGMFHSITFPTGTFVGTBWKTLVTILAIAPVLGIS 230
DB 100 GELMTRMT-DIDKISDIDGKIALIFQNMSTFSTIGLAVLWKGWKLVTLSLPLMAS 158
QY 231 AAIWAKLISFTDKELLAYAKAGAAVEVLAIRTVIAFGGQKELE----- 277
DB 159 AAACSRWVSIITSKELSAYSKAGAAVEVLSIRTVIARPAQKELOGRFELNITRYAMP 218
QY 278 -----RYKNLEEAKGIGIKAKITANISIGAFLLIYASTALAPWYGTSLV 323
DB 219 YFPQWMLSCVLXAFRYQYQNKDAKDFGIKKTASKVSLAAVYFNMGTGLAFWYGTSLI 278
QY 324 LSSE--YSIGOVLVFSPVLIGAFSIGQASPSIEAFANRGAAYEFKIIDKPSIDVS 381
DB 279 LNSPGYITIGTVLAVFVSYSYICGAANPHETAIARGAAPHIFQVIDKPSIDNS 338
QY 382 KSGKRPNINIGNLEFKNVHSYSPKREVKILGLNLKVQSGQTVALVNSGCGKSTVQL 441
DB 339 TAGKRPESIGTVEFKNVSEFNPSPRSIKILGLNLRIKSGETVALVGLNSGKSTVQL 398
QY 442 MORLYPTDGMVCIDGODIRITINVRHLREITGVSEPVLFATTTAENIRKRENVYMD 501
DB 399 LQRLYDPDGFIVDENDIRALNVRHRYKHIGVSOEPLVLTFTTISNNIKYGRDVTDE 458
QY 502 IEKAVKEANAAYDFIMKLPKFPDLVGERGAQLSGGQKORIAIARALVNNPKILIDEATS 561
DB 459 MEBAAREANAAYDFIMEFPMKFNMTLVGERGAQMSGGQKORIAIARALVNNPKILIDEATS 518
QY 562 ALDTESEAVYOVALLDKARKGRTTIYIAHLSTVRAADVIAGEDGVIVEKNHDELMEK 621
DB 519 ALDSEKSAVQALEKASGRTTIYVAHRLSTIRADLIVTLKQMLAEKGAHAEIMAKR 578
QY 622 GIYFKLVTVNQ-TRGNEIELENATGESKSESDALEM-SPKSDSSSLIKRSTRIRSHAPQ 679
DB 579 GLYLSIVMSQDIKKADEQMESMTYSTERKTNSLPHASKVSKISDPT----- 624
QY 680 QORKLGTREDLLENVPPVSFWRLKLNSTEMPEYVVGIFCAIINGLOPARSIFSRNIG 739
DB 625 -DKAESSTQSKESISLPEVSLKIILKLNKPWEPEVVLGTIASVLANGTVAPVSIIFAITT 683
QY 740 IFTRPDEPTKRONSMNMFSLVLYGIISFTTEPLOGTPEKAGEIILTKRIKRVFSSML 799
DB 684 MF-GNNDKTTLKHADEIYSMTIVILGIVCEFSYFMOGLFETRAGEIILTMRLRLHAFKRL 742
QY 800 RODVSWFDDPKFTGALTTRLANDAAOVKGAIGSRILAVITTONIANLGTGIIISLYQMOL 859

```

```

DB 743 YODIAMPKEKENSTGGTLTTLIAIDTAQIGATGSGTIGTLQNMATNMLSVIISPIYGM 802
QY 860 TLLLAIVPIIAIAGVEMKMLSGOALKDKKELGAGIATEALENRTVASTLRREKE 919
DB 803 TFLIISINPVLAVNGIMETAAATGFANKDKOELHAGKINTEALENIRTVISLTREKAFE 862
QY 920 VYKASLOVPYRNSLRKAHIFGVSEFSITQAMMYPSVAGCFRFGAYILVANFMFOVLAV 979
DB 863 QMYEEMLOTORHNRTSKKAOIIGSCYAFSHAFYAVAAAGRFAGYLQDARMTPEGMFIV 922
QY 960 FSAIVFGAMAYGVYSSFPADYAKAKVSAHVTIITEKSPILDSYSPHGLKPNLTESNVTF 1039
DB 923 FTALVGAALIGKTLVLAPEYSKAKSGAAHFLFALLEKPNIDISOSQGGKPPDCEGNLEF 982
QY 1040 NEVVNPTRPDIIVLOGLSLEVKKGOTLALVSGSGGKSTVOLLEREPDPLAGSVLD 1099
DB 983 REVSFETPCRPDVFILNGLSISTRGKTVAFVSSGCGKSTVQLQRLDIPVQGOVLFD 1042
QY 1100 GKELKHLNVOMLRAHLGIVSOEPLIFDCSIAENIAYGDNRSRVVSHBEIMOAKENIHFF 1159
DB 1043 GVDKAKELNVQWLRQIAIVPOEPLVFNCSIAENIAYGDNRSRVVPLDEIKEAANAANHSF 1102
QY 1160 IETLPKNTFVGPDKGTVLGGGQKORIAIARALVROPHILLIDEATSALDTESEKVVQEA 1219
DB 1103 IEGPEKNTNVOGLKGAQSGGQKORIAIARALLOKPKILLIDEATSALDNOSEKVVQHA 1162
QY 1220 LDKAREGRTCIIVAHRLSTIONMDLIVFPNGKVKKEGTHOOLAKGIVFSVSVQ 1277
DB 1163 LDKARTGRTCLIVTHRLSAIONMDLIVLHNGKIKKQGTGHELLRNRDIYFKLVANQS 1220

```

```

RESULT 11
US-09-873-409-4
; Sequence 4, Application US/09873409
; Patent No. US20020037522A1
; GENERAL INFORMATION:
; APPLICANT: Sayegh, Mohamed
; TITLE OF INVENTION: A Gene Encoding a Multidrug Resistance Human P-Glycoprotein
; FILE REFERENCE: 81994/268611
; CURRENT APPLICATION NUMBER: US/09/873.409
; CURRENT FILING DATE: 2001-06-05
; NUMBER OF SEQ ID NOS: 19
; SOFTWARE: PatentIn version 3.0
; SEQ ID NO 4
; LENGTH: 1058
; TYPE: PRT
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: No. US20020037522A1e
; LOCATION: (66)..(66)
; OTHER INFORMATION: Xaa at position 66 represents any L amino acid
US-09-873-409-4

```

```

Query Match          48.1%; Score 3112.5; DB 10; Length 1058;
Best Local Similarity 56.2%; Pred. No. 9.1e-231;
Matches 601; Conservative 201; Mismatches 220; Indels 47; Gaps 6;

QY 240 SFTDKELLAYAKAGAAVEVLAIRTVIAFGGQKELE----- 277
DB 4 SLTSKELSAYSKAGAAVEVLSIRTVIAFRAGKELEGRSLNLNITRYAMPYFPQWLSC 63
QY 278 -----RYKNLEEAKGIGIKAKITANISIGAFLLIYASTALAPWYGTSLVLSSE--YSI 330
DB 64 VLYKRVYTONLNDADDFGIKRTIASKVSIGAVYFNMGTGYLAFWYGTSLINDEPGYTI 123
QY 331 GOVLTVFESVLIGASQASPSIEAFANRGAAYEFKIIDNKPDSIDYSKSGKRPDI 390
DB 124 GTVLAVFVSYSYICGAANPHETAIARGAAPHIFQVIDKPSIDNSYAGKRPESI 183
QY 391 KGNLEFKNVHSYSPKREVKILGLNLKVQSGQTVALVNSGCGKSTTVQIMQRLYDPTD 450

```

[illegible]

```

1 TITLE OF INVENTION: Resistance to Macrocyclic Lactone Compounds
2
3 NUMBER OF SEQUENCES: 12
4
5 CORRESPONDENCE ADDRESS:
6 ADDRESSEE: American Home Products Corporation
7 STREET: One Campus Drive
8 CITY: Parsippany
9 STATE: New Jersey
10 COUNTRY: United States
11 ZIP: 07054
12
13 COMPUTER READABLE FORM:
14 MEDIUM TYPE: Floppy disk
15 COMPUTER: IBM PC compatible
16 OPERATING SYSTEM: PC-DOS/MS-DOS
17 SOFTWARE: PatentIn Release #1.0, Version #1.30
18
19 CURRENT APPLICATION DATA:
20 APPLICATION NUMBER: US/09/749,340
21 FILING DATE:
22 CLASSIFICATION:
23
24 PRIOR APPLICATION DATA:
25 APPLICATION NUMBER: 09/067,676
26 FILING DATE:
27
28 ATTORNEY/AGENT INFORMATION:
29 NAME: Renda, Barbara L.
30 REGISTRATION NUMBER: 27,626
31 REFERENCE/DOCKET NUMBER: 33,333-01
32 TELECOMMUNICATION INFORMATION:
33 TELEPHONE: 973-683-2153
34
35 TELEFAX: 973-683-4109
36 INFORMATION FOR SEO ID NO: 6:
37 SEQUENCE CHARACTERISTICS:
38 LENGTH: 1275 amino acids
39 TYPE: amino acid
40 STRANDEDNESS: single
41 TOPOLOGY: linear
42 MOLECULE TYPE: protein
43
44 US-09-749-340-6
45
46
47
48
49
50
51
52
53
54
55
56
57
58
59
60
61
62
63
64
65
66
67
68
69
70
71
72
73
74
75
76
77
78
79
80
81
82
83
84
85
86
87
88
89
90
91
92
93
94
95
96
97
98
99
100
101
102
103
104
105
106
107
108
109
110
111
112
113
114
115
116
117
118
119
120
121
122
123
124
125
126
127
128
129
130
131
132
133
134
135
136
137
138
139
140
141
142
143
144
145
146
147
148
149
150
151
152
153
154
155
156
157
158
159
160
161
162
163
164
165
166
167
168
169
170
171
172
173
174
175
176
177
178
179
180
181
182
183
184
185
186
187
188
189
190
191
192
193
194
195
196
197
198
199
200
201
202
203
204
205
206
207
208
209
210
211
212
213
214
215
216
217
218
219
220
221
222
223
224
225
226
227
228
229
230
231
232
233
234
235
236
237
238
239
240
241
242
243
244
245
246
247
248
249
250
251
252
253
254
255
256
257
258
259
260
261
262
263
264
265
266
267
268
269
270
271
272
273
274
275
276
277
278
279
280
281
282
283
284
285
286
287
288
289
290
291
292
293
294
295
296
297
298
299
300
301
302
303
304
305
306
307
308
309
310
311
312
313
314
315
316
317
318
319
320
321
322
323
324
325
326
327
328
329
330
331
332
333
334
335
336
337
338
339
340
341
342
343
344
345
346
347
348
349
350
351
352
353
354
355
356
357
358
359
360
361
362
363
364
365
366
367
368
369
370
371
372
373
374
375
376
377
378
379
380
381
382
383
384
385
386
387
388
389
390
391
392
393
394
395
396
397
398
399
400
401
402
403
404
405
406
407
408
409
410
411
412
413
414
415
416
417
418
419
420
421
422
423
424
425
426
427
428
429
430
431
432
433
434
435
436
437
438
439
440
441
442
443
444
445
446
447
448
449
450
451
452
453
454
455
456
457
458
459
460
461
462
463
464
465
466
467
468
469
470
471
472
473
474
475
476
477
478
479
480
481
482
483
484
485
486
487
488
489
490
491
492
493
494
495
496
497
498
499
500
501
502
503
504
505
506
507
508
509
510
511
512
513
514
515
516
517
518
519
520
521
522
523
524
525
526
527
528
529
530
531
532
533
534
535
536
537
538
539
540
541
542
543
544
545
546
547
548
549
550
551
552
553
554
555
556
557
558
559
560
561
562
563
564
565
566
567
568
569
570
571
572
573
574
575
576
577
578
579
580
581
582
583
584
585
586
587
588
589
590
591
592
593
594
595
596
597
598
599
600
601
602
603
604
605
606
607
608
609
610
611
612
613
614
615
616
617
618
619
620
621
622
623
624
625
626
627
628
629
630
631
632
633
634
635
636
637
638
639
640
641
642
643
644
645
646
647
648
649
650
651
652
653
654
655
656
657
658
659
660
661
662
663
664
665
666
667
668
669
670
671
672
673
674
675
676
677
678
679
680
681
682
683
684
685
686
687
688
689
690
691
692
693
694
695
696
697
698
699
700
701
702
703
704
705
706
707
708
709
710
711
712
713
714
715
716
717
718
719
720
721
722
723
724
725
726
727
728
729
730
731
732
733
734
735
736
737
738
739
740
741
742
743
744
745
746
747
748
749
750
751
752
753
754
755
756
757
758
759
760
761
762
763
764
765
766
767
768
769
770
771
772
773
774
775
776
777
778
779
780
781
782
783
784
785
786
787
788
789
790
791
792
793
794
795
796
797
798
799
800
801
802
803
804
805
806
807
808
809
810
811
812
813
814
815
816
817
818
819
820
821
822
823
824
825
826
827
828
829
830
831
832
833
834
835
836
837
838
839
840
841
842
843
844
845
846
847
848
849
850
851
852
853
854
855
856
857
858
859
860
861
862
863
864
865
866
867
868
869
870
871
872
873
874
875
876
877
878
879
880
881
882
883
884
885
886
887
888
889
890
891
892
893
894
895
896
897
898
899
900
901
902
903
904
905
906
907
908
909
910
911
912
913
914
915
916
917
918
919
920
921
922
923
924
925
926
927
928
929
930
931
932
933
934
935
936
937
938
939
940
941
942
943
944
945
946
947
948
949
950
951
952
953
954
955
956
957
958
959
960
961
962
963
964
965
966
967
968
969
970
971
972
973
974
975
976
```


Db 599 LIDEATSDLTSEKGVQALERRAEGRTTIIAHLSTIKTAHNTIVLVNKGIAEQGH 658
QY 615 DELMKRGYIKFLVTMQTGNIEI---LENNAGESKESSEDALEMSPKDSS-----L 664
Db 659 DELVDGGAIRKIVEAQRIKNECKEADALDADAEDLTNDIAKIKIYASSASSDLDOKPTT 718
QY 665 IKRRSTRSRSHADQODRKIGTKEDINENVPYVFWRIIK---LNSTWPTFVVGIFCA 720
Db 719 IDRTGTHKSVSSAIIKSRP-----PETTPKYSIWLTLKFAVSFRRPEIPYMLIGLVFS 771
QY 721 IINGGLOPASITFSRIGITFTRDEDEPERK-RONSMMFSVFLVLGIISFTIFFLOGFTF 779
Db 772 VLAGGGOPQAVLYAKAISTSLPESQYSLRHADPWSLMEPVYVGIIOFTIQTSTGAPF 831
QY 780 GKAGEILTKRLRYVFRSMRLQDVSWEDPKNTTGALTTRLANDAAQVKAIGSLAVIT 839
Db 832 AVCSERILRRARSTAFRTILRODIAFDKRENSGTALTSFLSTETKHLGSGSVTLGITL 891
QY 840 ONIANIGTGIIISLYGMQTLTLLAIVPIIALAGVEMKMSGALKKKELEGAKTA 899
Db 892 MTSTTLGAALIIIALIGMKLALVCISVYPVLLACGFYREYMLAQFQSRSLAYEGSANRA 951
QY 900 TEATENRTVVSILTRQCKFEYMAQSLQVRYRNSLRKAHIFGVSFILTOAMTFSYAGCF 959
Db 952 CEATSSIRTYASLTFREDVWEIYHAOLDAGRTSLISVLNSSLVASSQALVEFCVALGF 1011
QY 960 RFQAYIVANEFMNEQDVLVFSALVFGAMAVGOVSFAPDYAKAKYSAHVIMIEKSP 1019
Db 1012 WYGTTLGHHHEYDIFRFVCFSEILFGAQSAGVYFSPAPMGAKNAAMAFRRILFRKPO 1071
QY 1020 IDSSPGILKPNTEGNAVTFNEVYVNPTRPDIIPVLOGLSLEVKKGOTLLAVSSGCGKS 1079
Db 1072 IDNMSSEBEKLEIYVEGIEFRNNAHFRTPTPEQPVLRKGLDLYTKPGYVALVPSGCGKS 1131
QY 1080 TVVOLLERFYDPLAGSVLIDKEIKHLNVOLMLRHILGIVSQEPLTDCSIAENIANGDS 1139
Db 1132 TTLLEEFYDALIGSLIVDCKDISKININSYRSFLSLVQSEPLLYGXTIKENILGIVE 1191
QY 1140 RVVSHHELMQAAKANHHFETLPEKYNRYGDKQTQSGGOKORAIARALVROPHIL 1199
Db 1192 DDVVEEFLIKCKDANIYDFLMSLPESFNIVVSGKMSLGGQORVAINAKALNDPKIL 1251
QY 1200 LLDERTSALDTESEKVVQALDKAREGRTCIIVIAHRLSTIQNDLIVVFOGKVEHGT 1259
Db 1252 LLDERTSALDSESEKVVQALDAARGRITIIAHHRLSTIQKADVIIVFPGKIVESGTH 1311
QY 1260 QQLLAQKGIYSWVSQAGAK 1280
Db 1312 SELVQKGRYYELVLOSIGK 1332

RESULT 14

US-09-873-409-2
; Sequence 2, Application US/09873409
; Patent No. US20020037522A1
; GENERAL INFORMATION:
; APPLICANT: Frank, Markus
; TITLE OF INVENTION: A Gene Encoding a Multidrug Resistance Human P-Glycoprotein
; FILE REFERENCE: 81994/268611
; CURRENT APPLICATION NUMBER: US/09/873,409
; CURRENT FILING DATE: 2001-06-05
; NUMBER OF SEQ ID NOS: 19
; SOFTWARE: PatentIn version 3.0
; SEQ ID NO 2
; LENGTH: 812
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-873-409-2

Query Match 37.7%; Score 2442; DB 10; Length 812;
Best Local Similarity 56.8%; Pred. No. 1,9e-179;

Matches 469; Conservative 160; Mismatches 178; Indels 18; Gaps 4;
QY 455 IDGODIRINRHLREITGVVSQEPVLFATIIANINRYGRNVMDIEKAVKANAAYDE 514
Db 2 VDENDIRALNRYRHDHIGVVSQEPVLEGTISNNIKYGRDVTJDEMERARANAYDF 61
QY 515 IMKLPNKFDTLVGERGASQSGOKORIALARALVNRKILLDDEATSDLTSESAVVOA 574
Db 62 IMEPNKNNTLVGERGASQSGOKORIALARALVNRKILLDDEATSDLTSESAVVOA 121
QY 575 LDKARGRFTTIVAHRLSTVRNADVIAGFDGVYVEKGNDELKKEKGIYFKLVTMQ-TR 633
Db 122 LEKASKGRTTIVAHRLSTIRASDLIVTLKDMGLAEKGAHLEAKKGLVYSLVMSDID 181
QY 634 GNEITLENATGESSEDALEM-SPKDGSSILKRRSTRSHAPQODRKIGTKEDLINE 692
Db 182 KADEMESMTSTETKTSLSPLHSYKTSKSPF-----DKAESSTQKEI 226
QY 693 NVPYVFWRIILKNTSTEMPVYVGFCAIINGGLOPASITFSRIIGITFTRDEDEPERKQ 752
Db 227 SLPEVSLKIKTLKRNPEWPFVVLGTIASVNGTVHPVFSIIFAKIITWF-GNNDKTYLKH 285
QY 753 NSNMFVLEFLVIGITISFTIFFLOGFTFGKAGEILTKRLRYVFRSMRLQDVSWFDDPKNT 812
Db 286 DAEIYSMIFVLIVGICFVSFEMQGLFGRAGELTMRRLHAFKAMLYODIANTWDEKENS 345
QY 813 TGAULTRLANDAAQVKAIGSLRAVITTONIANLGTIIISLYGMQTLTLLAIVPIIAI 872
Db 346 TGGITTLIALIDIAQIQAGATGSRIGVLTQNTMGLSVIISFYGEMFPLISLAPVLAV 405
QY 873 AGVNEKMLSGQALKDKKELEGAKIATEAIENFTVVSILTRQCKFEYMAQSLQVRYR 932
Db 406 TGMIEETAMTGEANDKDKELHAGKIAATEALEINIFTIVSLTRKAEBOYEMBLDQHRN 465
QY 933 SLRKAHIEGVSFSTIQAMAFYFACFRGAYLVANEFQDVLVFSALVFGAMAVGO 992
Db 466 TSKRAQIDISCTAFSIAFIFYFAAGFRGATLIQAGNTPGCMIVFTALAYGMAIGK 525
QY 993 VVSFAPDVAKAVYSAHVIMIEKSPIDISYSPHGLKPTLLEGNAVTFNEVYVNPTRPDI 1052
Db 526 TLYLAEYSKASGAHNLPALEKRPNDISRSQDEKKPDTGCGNLEFREFVSFFYCRPDV 585
QY 1053 PVIQGISLEVKKGOTIALVSGSGCKSTVYOLLEBFYDPLAGSVLIDKEIKHLNVQMLR 1112
Db 586 FILRGSLSTIERKTYVAFVSGSGCKSTSVOLLRKYDVOGOVLPDGYDAKELANVQMLR 645
QY 1113 AHHGIYSOEPILFDCSIAENIAYGDSRVVSHHEIMQAAKANHHFETLPEKYNRYG 1172
Db 646 SQIAIYPOEPVLFNCISIAENIAYGDSRVVPLDEIKKAANMANIHSFIEGLPEKYNTOVG 705
QY 1173 DKGTQSGGOKORIALARALVROPHILLDDEATSDLTSEKVVQALDKAREGRTCIIV 1232
Db 706 LKGAQLSGCKOKRIARALRLOKPKILLDDEATSDLDNSEKVVQALDALKARTGRTCLAV 765
QY 1233 AHHRLSTIONADLIVFONGKVEKHTHOOILAOKGIYFMSVQOA 1277
Db 766 THRLSATIONADLIVLVHNGKIKEGTHOELRNRODLYFRLVNAQS 810

RESULT 15

US-10-101-388-3
; Sequence 3, Application US/10101388
; Patent No. US20020162142A1
; GENERAL INFORMATION:
; APPLICANT: Johanl, Gurmukh S
; APPLICANT: Multani, Dilbag S
; APPLICANT: Briggs, Steven P
; TITLE OF INVENTION: GENES AND METHODS FOR MANIPULATION OF GROWTH
; FILE REFERENCE: 5718-81 (035718/205794)
; CURRENT APPLICATION NUMBER: US/10/101,388
; CURRENT FILING DATE: 2002-03-19
; PRIOR APPLICATION NUMBER: 09/711,562
; PRIOR FILING DATE: 2000-11-13


```

; PRIOR APPLICATION NUMBER: 60/164,886
; PRIOR FILING DATE: 1999-11-12
; NUMBER OF SEQ ID NOS: 3
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 3
; LENGTH: 1394
; TYPE: PRN
; ORGANISM: Zea mays
US-10-101-388-3

```

```

Query Match      30.1%: Score 1950; DB 9; Length 1394;
Best Local Similarity 36.1%: Pred. No. 2,6e-141;
Matches 479; Conservative 235; Mismatches 468; Indels 146; Gaps 28;

```

```

QY 25 NEKKERKPTVSTFAMFRYSNMIDRLYMLVGTMAIITHAALPLMLLVGNMDSFANAIGI 84
DB 113 NDSKKPTPPAALRDLPFRFADGDCAIMLIGTIGALVHCISLPVLRFPADLVDSFGSHAD 172
QY 85 SRKKTTPVLIINSTNNTOHFINLLEEMKTYAAYYSIGAGVLVAAYIOVSEWCLAAGR 144
DB 173 DDPFMVRLVVK-----YAFYFLVVGAAIWAASSMAEISCMMTGER 212
QY 145 QLIKIKOFFHAIMROEIGMFDVHDVGLNTRLDDVSKIEGDKTGMFHSIAIFFT 204
DB 213 QSTRMIRIRLDALAQDVSFPDT-----DVRASDVYIAINADAVVVOGR-HQPETGQP 264
QY 205 GTVGTGRCWKLTVLILAIISPVLSAAMAKILSFTDKELLAYAKAGAV-----AEE 258
DB 265 HPLHGLKNG-----RLRGVHGRVAAAG-----AGHARGAANRRHRAER 304
QY 259 VLAALRTYIANG-----GQKKELEYKNLBEAK-----GIGIKKAITANT 299
DB 305 RRAARQALVPOQRALGRRHRGAGARADTVAGVRMGARGAGLLGGAGGADRLPQRL 364
QY 300 SIGAA-----FLIYAYALAFWYGTSLVLSSEYSIGOVLFYFVSLGARSIGQASP 352
DB 365 RQGARPRRHLLHRLLLLPALVLRPPRARAHORRRARHNTM---FSVMIGG-GPRQSA 420
QY 353 STEAFANRGAAYELEFKTIDNKPSIDSYSKSGHKPDNIKGNLEFKVNFESYPSRKEVKIL 412
DB 421 SMAAPAKAKVAARKIFRIIDHRPGIS--SRDGAPEBSVTGVRMGRGVDFAYPSRPDPIL 478
QY 413 KGLNLKYOSGQTVALVGNSSGCKSTVOLMORLYDPTDGVWCIDGODIRTIANRHLREIT 472
DB 479 RGFSLSVAGKTIALVSSSGSKSTVSLIERFYDPSAGQILLDGHLRSLLEMLRQI 538
QY 473 GVYSGOEPLVFTTAENIRYGR--ENVTMDEIKKAVKEANAYDFIMKLPNKEDTLVGERG 530
DB 539 GLVSGOEPLFATSIRENILLGRDOSATLAEMEEAARVANAHSFIIKLDPGYDTVOYGERG 598
QY 531 AOLSGGOKORAIARALVVRPKIILLDEATSAIDTESAVVOVALDKARKGRTTI-VIAH 589
DB 599 LQLSGGOKORAIARALKKIPAILLDEATSAIDSESEKLYOEDALDRFMKGRTTLCDRAT 658
QY 590 RLSTVRNADYIAGFDVIVIEKGNHDEIMK--EKGIYFKLVMTQTRGNETIELENATGESK 647
DB 659 GCPSPAKADYVAVVLOGGAVSEMAMHDEIMAKGNGTYAKLIMQEOGAHEALVNA----- 713
QY 648 SEDALMSPKDSGSSLIKRRS-----TRR-----STHAPOGDRKLGTR 688
DB 714 RRSARPSARSNSVSPIMTRNSYGRSPSRISDFTLSIDPHNNHRTMADKQ 773
QY 689 DLNENVPVPSFMRILKINSTEMPYFVVGICAIINGLOPAFSIIFSRIGTIGFTDEDEPE 748
DB 774 -LAFRAGASSFLRLARNNSPEMAYALAGISGVCCSFSAIFAYILISALSVYYA--PD 829
QY 749 TKRONSMMSVLFLVGLIISFTTF--LOGFTGKAGELITLRLRYMFRSMLRQDVSMF 806
DB 830 PRYMKREIAYCYLLIGMSSAALLENTYOHVFWDFVGENLTKRVREKMFVAVERNEIMAF 889
QY 807 DDPKNTTGALTTRLANDAQAQVGAIGSRILAVTTOMIANLGIIISLIYQOLTLILLAI 866
DB 890 DADENMSAKVTATLALDAQONRSALGDRISIVONSAIMLVACTAGFVLQWRALVLILAV 949

```

```

QY 867 VPIIATAGVE---MKMLSGOALKDKKELEG----AGIATEAIEINFRTVVSILTRQKFE 919
DB 950 FPLVGAATVLOKKFMKGFSG-----DLEAHARATQIAGAVANLNTVAFAAERKIT 1002
QY 920 YMYAOSLOVYRNSLRKAHIFGVSVSIQAMMYEYIACCFRGAIVLANEFMNFQDVLV 979
DB 1003 GLEFANLRQPLRRCFWKGOIAGSGYGAOFLLYASVALGLMYAAMLVKGVSDFSRITRV 1062
QY 980 FSAIYEGAMAVOVSSFAPDYAKAKVSAAHVIMIERKSPILDSYSPGLK-----PNTLEG 1035
DB 1063 FMYLWVSANGAAETLILAPDFIKGRARMSVFETIDRKTVE--PHVDAAAPVDPDGA 1119
QY 1036 NTTFENVYNTTRPDIPYLOGLSLEVKKGGOTLALVSSGGCKSTVQVLERFYDPLAGS 1095
DB 1120 KYELKHVDLPSPRPDIQVFRDLSIRARAKTLALVGPSSGCKSSVALVQRFYKPTSGR 1179
QY 1096 VLIDGKEIKHLNVOMLRAHLGIVSOEPIIFDCSTAEINAYGDNRRVVSHEEIMOAKEAN 1155
DB 1180 VLLDGKDVAKYKURLRKRNVAVVPOEPPLFAASIHENIAYGREG--ATEAEVEAAAOAN 1237
QY 1156 IHHFTETLPEKYNTFVGDGKTQLSGGOKORAIARALVROPHILLDEATSAIDTSEKXV 1215
DB 1238 AHRFTALPEGYRTQVGERGVQLSGGOKORAIARALVKKOALVLLDEATSAIDASEERC 1297
QY 1216 VOEALDKAREGTCIVIAHRLSTIQNADLIYFONGKVKKEGHQOLLAO--KGIYFSNV 1273
DB 1298 VOEALERAGSGRTTIVVAHRLATYGAHTTIAVIDGKVAEGSSHSHLKHHPGCVYARM 1357
QY 1274 SVOAGAKR 1281
DB 1358 QLAADGR 1365

```

Search completed: December 9, 2002, 17:12:06
Job time : 151 secs

GenCore version 5.1.3
Copyright (c) 1993 - 2002 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: December 9, 2002, 16:20:53 ; Search time 639 Seconds
(without alignments)
192.720 Million cell updates/sec

Title: US-09-672-725c-2
Perfect score: 6477
Sequence: 1 MPPEGGRKSAEKNFWKMKK.....LLAQKGIYSVMVYQAGAKR 1281

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 283224 seqs, 96134422 residues

Total number of hits satisfying chosen parameters: 283224

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database :

1: p1r1:*
2: p1r2:*
3: p1r3:*
4: p1r4:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	5861	90.5	1280	1 DVH01	multidrug resistan
2	5651.5	87.3	1276	1 DVH1C	multidrug resistan
3	5633	87.0	1276	1 A34786	multidrug resistan
4	5296	81.8	1276	1 DVMS1	multidrug resistan
5	5226.5	80.7	1277	2 JH0502	multidrug resistan
6	4990	77.0	1104	1 DVMS1A	multidrug resistan
7	4905.5	75.7	1279	1 DVH03	multidrug resistan
8	4856	75.0	1276	1 DVMS2	multidrug resistan
9	4818.5	74.4	1281	2 I48123	multidrug resistan
10	4803	74.2	1278	2 S41646	multidrug resistan
11	4422	68.3	1287	2 S53692	multidrug resistan
12	3209.5	49.6	1321	2 T42228	multidrug resistan
13	3205.5	49.5	1321	2 T42228	multidrug resistan
14	2857	44.1	1294	2 T42842	multidrug resistan
15	2819	43.5	1289	2 D19982	multidrug resistan
16	2818.5	43.5	1275	2 T31073	multidrug resistan
17	2741	42.3	1321	2 T23476	multidrug resistan
18	2739	42.3	1321	2 S27337	multidrug resistan
19	2674.5	41.3	655	1 DVH2C	multidrug resistan
20	2651	40.9	1283	1 A47377	multidrug resistan
21	2602.5	40.1	1286	2 T02187	multidrug resistan
22	2596	40.1	1292	2 T48007	multidrug resistan
23	2589.5	39.6	1378	2 E86155	multidrug resistan
24	2564.5	39.6	1302	2 A41249	multidrug resistan
25	2521	38.9	1229	2 D85023	multidrug resistan
26	2518	38.9	1229	2 T52319	multidrug resistan
27	2518	38.9	1230	2 E85023	multidrug resistan
28	2496.5	38.5	1302	2 B41249	multidrug resistan
29	2473	38.2	1229	2 F86155	multidrug resistan

30	2441	37.7	1408	2 T43261	multidrug resistan
31	2437.5	37.6	1323	2 H85202	multidrug resistan
32	2431	37.5	1286	2 A42150	multidrug resistan
33	2423	37.4	1310	2 S30328	multidrug resistan
34	2406	37.1	1266	2 T22090	multidrug resistan
35	2398	37.0	1268	2 T22094	multidrug resistan
36	2380	36.7	1302	2 S30327	multidrug resistan
37	2345	36.2	1254	2 S27338	multidrug resistan
38	2330.5	36.0	1222	2 T14805	multidrug resistan
39	2307	35.6	1233	2 T04251	multidrug resistan
40	2297.5	35.5	1245	2 G86404	multidrug resistan
41	2289	35.3	1254	2 T30855	multidrug resistan
42	2287	35.3	1307	2 T30882	multidrug resistan
43	2277	35.2	1362	2 T41534	multidrug resistan
44	2269.5	35.0	1318	2 T21266	multidrug resistan
45	2246	34.7	1247	2 F86405	multidrug resistan

ALIGNMENTS

RESULT 1
DVH01
multidrug resistance protein 1 - human
N:Alternate names: P-glycoprotein 1
C:Species: Homo sapiens (man)
C>Date: 31-Dec-1990 #sequence_revision 18-Aug-1995 #text_change 19-Jan-2001
C:Accession: A34914; PS0162; S15500; A25059; S43838; I52238; I65204
R:Chen, C.; Clark, D.; Ueda, K.; Pastan, I.; Gottesman, M.M.; Roninson, I.B.
J. Biol. Chem. 265, 506-514, 1990
A:Title: Genomic organization of the human multidrug resistance (MDR1) gene and origin
A:Reference number: A34914; MUID:9004448; PMID:1967175
A:Accession: A34914
A:Molecule type: DNA
A:Residues: 1-1280 <CHE>
A:Cross-references: GB:M29447; GB:J05168; NID:q187496; PIDN:AAAS9576.1; PID:g386862
R:Kioka, N.; Yamano, Y.; Komano, T.; Ueda, K.
submitted to JRPID, April 1991
A:Reference number: PS0162
A:Accession: PS0162
A:Molecule type: DNA
A:Residues: 1-22 <KIO>
R:Kioka, N.; Yamano, Y.; Komano, T.; Ueda, K.
submitted to the EMBL data library, April 1991
A:Description: Transcriptional regulation of multidrug resistance gene (MDR1) express
A:Reference number: S15500
A:Accession: S15500
A:Molecule type: mRNA
A:Residues: 1-22, 'R' <K12>
A:Cross-references: EMBL:X58723; NID:g34522; PIDN:CAAA1558.1; PID:g34523
R:Chen, C.; Chin, J.E.; Ueda, K.; Clark, D.P.; Pastan, I.; Gottesman, M.M.; Roninson, I.B.
Cell 47, 381-389, 1986
A:Title: Internal duplication and homology with bacterial transport proteins in the m
A:Reference number: A25059; MUID:87028230; PMID:2876781
A:Accession: A25059
A:Molecule type: DNA
A:Residues: 1-184, 'V', 186-1280 <CH2>
A:Cross-references: GB:M14758; NID:g187468; PIDN:AAAS9575.1; PID:g307180
R:Chambers, T.C.; Pohll, J.; Glass, D.B.; Kuo, J.F.
Biochem. J. 299, 309-315, 1994
A:Title: Phosphorylation by protein kinase C and cyclic AMP-dependent protein kinase
A:Reference number: S43838; MUID:94220047; PMID:7909431
A:Accession: S43838
A:Molecule type: protein
A:Residues: 656-689 <CHA>
R:Geckeler, V.; Weger, S.; Probst, H.
Biochem. Biophys. Res. Commun. 169, 796-802, 1990
A:Title: mdrl/P-glycoprotein gene segments analyzed from various human leukemic cell
A:Accession: I52238; MUID:90290529; PMID:1972623
A:Status: translated from GB/EMBL/DBJ
A:Molecule type: DNA
A:Residues: 178-215 <RES>

A:Cross-references: GB:M37724; NID:g183537; PIDN:AAA8047.1; PID:g553314
A:Accession: 165204
A:Status: translated from GB/EMBL/DBJ
A:Molecule type: DNA
A:Residues: 800-856 <RR2>
A:Cross-references: GB:M37725; NID:g183538; PIDN:AAA8048.1; PID:g553315
A:Comment: This is an integral membrane protein overproduced in multidrug-resistant cell
structurally and functionally unrelated lipophilic anticancer drugs.
C:Genetics:
A:Gene: GDB:PGY1:MDR1
A:Cross-references: GDB:120712; OMIM:171050
A:Map position: 7q21-7q21
A:Superfamily: multidrug resistance protein; ATP-binding cassette homology
C:Keywords: ATP; duplication; glycoprotein; nucleotide binding; P-loop; phosphoprotein;
F:1-638,653-1280/Region: duplication
F:49-350/Domain: hydrophobic <HB1>
F:351-637/Domain: hydrophobic <HB1>
F:410-604/Domain: ATP-binding cassette homology <ABC1>
F:427-434/Region: ATP-binding nucleotide-binding motif A (P-loop)
F:551-555/Region: nucleotide-binding motif B
F:638-708/Domain: linker <LIN>
F:709-993/Domain: hydrophobic <HB2>
F:994-1280/Domain: hydrophobic <HL2>
F:1053-1249/Domain: ATP-binding cassette homology <ABC2>
F:1070-1077/Region: nucleotide-binding motif A (P-loop)
F:1196-1200/Region: nucleotide-binding motif B
F:91-94,99/Binding site: carbohydrate (Asn) (covalent) #status predicted
F:433/Binding site: ATP (Lys) #status predicted
F:661,667,671/Binding site: phosphate (Ser) (covalent) (by CAMP-dependent kinase) #statu
F:1076/Binding site: ATP (Lys) #status predicted

Query Match 90.5%; Score 5861; DB 1; Length 1280;

Best Local Similarity 90.7%; Pred. No. 0;
Matches 1163; Conservative 55; Mismatches 60; Indels 4; Gaps 4;

QY 1 MPEGGKRGSA-EKNEFWKSKSKNEKEKKPYSTAMRYSNWLDRLMLVGTMAAI 59
Db 1 MLEGGKRGCAKKKNEFKLNKSEK-DKEKKPYSVSMRYSNWLKLMVGTILAI 59
QY 60 IHCAALPLMLVGNMTDSFANAGISRNKTFPIINESITNTQHFINHLEEMTYAVY 119
Db 60 IHGAGPLMLVGEWTDIFANAG-NLEDMENITNSINDTGFPM-LEEDMTYAVY 117
QY 120 YSGIGAVLVAAYIYVSEWCLAGROILKIROFHAIKROIGEDVDHDELTRILD 179
Db 118 YSGIGAVLVAAYIYVSEWCLAGROILKIROFHAIKROIGEDVDHDELTRILD 177
QY 180 DVSKINGIDDKIGMFHSAIFETFGIYFTGRMKLTFLVLAISPVLAGLSAAIWKILS 239
Db 178 DVSKINGIDDKIGMFHSAIFETFGIYFTGRMKLTFLVLAISPVLAGLSAAVAKILS 237
QY 240 SFTDKELLAAKAGAAVEVLAIRTVLAFGOKKELEKYNKLEBAKGIIGKAITANI 299
Db 238 SFTDKELLAAKAGAAVEVLAIRTVLAFGOKKELEKYNKLEBAKGIIGKAITANI 297
QY 300 SIGAFLIYASVLAFTWGTSTVLSSEYSGOVLTFVFSVLGAFSISQASPSIEAFAN 359
Db 298 SIGAFLIYASVLAFTWGTSTVLSSEYSGOVLTFVFSVLGAFSISQASPSIEAFAN 357
QY 360 ARGAAVEIFKIIDNKPISIDSYSGKRPDIKNLEKKNVHESYPSRKEVYIKGLNLKV 419
Db 358 ARGAAVEIFKIIDNKPISIDSYSGKRPDIKNLEKKNVHESYPSRKEVYIKGLNLKV 417
QY 420 QSGQVALVNGSGCKSTTVQLMORLDPDWDGVCIDGODIRFINVRLHLEITGVVSQEP 479
Db 418 QSGQVALVNGSGCKSTTVQLMORLDPDWDGVCIDGODIRFINVRLHLEITGVVSQEP 477
QY 480 VLEATTIENIRYGRRENTMDIEKAKVKAANAADYFIKLNKFDVLVGERGALSGGQXQ 539
Db 478 VLEATTIENIRYGRRENTMDIEKAKVKAANAADYFIKLNKFDVLVGERGALSGGQXQ 537
QY 540 RIAIARALVRNPKIILLDEATSDLTSEANVVOVALDKARKGRITTVIAHRLSTVRNAV 599

Db 538 RIAIARALVRNPKIILLDEATSDLTSEANVVOVALDKARKGRITTVIAHRLSTVRNAV 597
QY 600 IAFEDGVIEYKGNHDELKKEGIFYKLTVMOTRGNEIELENAGSESDALEMRPKD 659
Db 598 IAFEDGVIEYKGNHDELKKEGIFYKLTVMOTRGNEIELENAGSESDALEMRPKD 657
QY 660 SGSSLIKRNSTRSHAPQOQRKIGTKEDLNENVPVSPFIRILKINSTENPYPVVGIFC 719
Db 658 SRSLSIKRSTRSVRSVQAOQRKIGTKEDLNENVPVSPFIRILKINSTENPYPVVGIFC 717
QY 720 AIIINGGQPAFSTIFSRIGITFRDDEPETKQNSNMSVLFVLGITSITFTFLOGTFE 779
Db 718 AIIINGGQPAFSTIFSRIGITFRDDEPETKQNSNMSVLFVLGITSITFTFLOGTFE 777
QY 780 GKAGEILTKRLRWVFSMLRODVSFDPDPKNTGALTTRLANDAAQVGAIGSRNAVIT 839
Db 778 GKAGEILTKRLRWVFSMLRODVSFDPDPKNTGALTTRLANDAAQVGAIGSRNAVIT 837
QY 840 QNIANLGTIIISITYGQMLTLAIVPILTAAGVEMKMSGQALDKKLEGAQKTA 899
Db 838 QNIANLGTIIISITYGQMLTLAIVPILTAAGVEMKMSGQALDKKLEGAQKTA 897
QY 900 TEATENFRTVSLTRKQEFYMTAQSLQVYRYSIKRAHIFGVFSITQAMVYFVAGCF 959
Db 898 TEATENFRTVSLTRKQEFYMTAQSLQVYRYSIKRAHIFGVFSITQAMVYFVAGCF 957
QY 960 REGATLVANEMNODVLYVSAIVFGAMAVGVVSFAPDYARAKVSAHVIMIEKSP 1019
Db 958 REGATLVANEMNODVLYVSAIVFGAMAVGVVSFAPDYARAKVSAHVIMIEKSP 1017
QY 1020 IDSYPHGLKPNTEGNTVENNVFNTPTRPDIYVLOGLSLEFKGGTALVSSGCGKS 1079
Db 1018 IDSYPHGLKPNTEGNTVENNVFNTPTRPDIYVLOGLSLEFKGGTALVSSGCGKS 1077
QY 1080 TYVOLLEPRYPDLASVYLIDGKEIKHLNVOMLRAHLIVSOEPLTFDCSIAENIAYGDS 1139
Db 1078 TYVOLLEPRYPDLASVYLIDGKEIKHLNVOMLRAHLIVSOEPLTFDCSIAENIAYGDS 1137
QY 1140 RVVSHHEIQAQKAEINHHFETLEPEKYNTRYGDKGTOLSGGQORAIARALVROPHIL 1199
Db 1138 RVVSHHEIQAQKAEINHHFETLEPEKYNTRYGDKGTOLSGGQORAIARALVROPHIL 1197
QY 1200 LIDEATSDLTSEANVVOVALDKARKGRITTVIAHRLSTIQNADLIYVFGNGVKEGTH 1259
Db 1198 LIDEATSDLTSEANVVOVALDKARKGRITTVIAHRLSTIQNADLIYVFGNGVKEGTH 1257
QY 1260 QOLLAQGIYFSWVSVOAGAKR 1281
Db 1258 QOLLAQGIYFSWVSVOAGAKR 1279

RESULT 2

DVH1C
Multidrug resistance protein 1 - Chinese hamster

N:Alternate names: P-glycoprotein pgp1

C:Species: Cricetus griseus (Chinese hamster)

C:Date: 31-Dec-1990 #sequence revision 30-Sep-1992 #text change 19-Jan-2001

C:Accession: A38696; C38696; B38696; A27126; S33768; I52823

R:Devine, S.E.; Hussain, A.; Davide, J.P.; Melzer, P.W.

J. Biol. Chem. 266, 4545-4555, 1991

A:Title: Full length and alternatively spliced ppg-1 transcripts in multidrug-resista

A:Reference number: A38696; M01D:91154265; PMID:1671863

A:Accession: A38696

A:Molecule type: mRNA

A:Residues: 1-1276 <DEV>

A:Cross-references: GB:M59253; NID:g191154; PIDN:AAA37004.1; PID:g191155

A:Accession: C38696

A:Molecule type: mRNA

A:Residues: 108-1276

A:Cross-references: GB:M59254; NID:g191156; PIDN:AAA37005.1; PID:g191157

A:Experimental source: Clone ADX185

A:Accession: B38696

A:Molecule type: mRNA
A:Residues: 1-32,771-1276 <DE2>
A:Cross-references: GB:M59252; NID:g191152; PIDN:AAA37003.1; PID:g191153
A:Experimental source: clone ADX124
R:Endicott, J.A.; Juranka, P.F.; Sarangi, F.; Gerlach, J.H.; Deuchars, K.L.; Ling, V.
Mol. Cell. Biol. 7, 4075-4081, 1987
A:Title: Simultaneous expression of two P-glycoprotein genes in drug-sensitive Chinese H
A:Accession: A27126; MUID:88122132; PMID:2893255
A:Status: not compared with conceptual translation
A:Molecule type: mRNA
A:Residues: 706-1276 <END>
A:Cross-references: GB:M17897; NID:g191158; PIDN:AAA37006.1; PID:g191159
R:Zastawny, R.L.; Ling, V.
Biochim. Biophys. Acta 1173, 303-313, 1993
A:Title: Structural and functional analysis of 5' flanking and intron 1 sequences of the
A:Reference number: S33768; MUID:93305724; PMID:8100449
A:Accession: S33768
A:Status: translation not shown
A:Molecule type: DNA
A:Residues: 1-21 <ZAS>
A:Cross-references: EMBL:L03286
R:Teeter, L.D.; Eckersberg, T.; Tsai, Y.; Kuo, M.T.
Cell Growth Differ. 2, 429-437, 1991
A:Title: Analysis of the Chinese hamster P-glycoprotein/multidrug resistance gene pgp1
A:Reference number: 152823; MUID:92088970; PMID:1661134
A:Accession: 152823
A:Status: translated from GB/EMBL/DBJ
A:Molecule type: DNA
A:Residues: 1-21 <RES>
A:Cross-references: GB:S81975; NID:g240862
C:Comment: This is an integral membrane protein overproduced in multidrug-resistant cell
structurally and functionally unrelated lipophilic antitumor drugs.
A:Gene: pgp1
C:Superfamily: multidrug resistance protein; ATP-binding cassette homology
C:Keywords: alternative splicing; ATP; duplication; glycoprotein; nucleotide binding; P-
F:407-601/Domain: ATP-binding cassette homology <ABC1>
F:424-431/Region: nucleotide-binding motif A (P-loop)
F:548-552/Region: nucleotide-binding motif B
F:1050-1246/Domain: ATP-binding cassette homology <ABC2>
F:1067-1074/Region: nucleotide-binding motif A (P-loop)
F:1193-1197/Region: nucleotide-binding motif B
F:87, 91, 96/Binding site: carbonylate (asn) (covalent) #status predicted
F:430/Binding site: ATP (lys) #status predicted
F:1073/Binding site: ATP (lys) #status predicted

Query Match 87.3% Score 5651.5; DB 1; Length 1276;
Best Local Similarity 86.8%; Pred. No. 0;
Matches 1113; Conservative 77; Mismatches 85; Indels 7; Gaps 3;

QY 1 MPEGGRGSAEKNFWMKKSKKNEKKEKPPVSTFAMFRYRNWLDRLYMVGMAAI 60
DB 1 MFEEDFSGRDKNFLEKMKRSK-K-EKKEKKPVSVFTMRVAYGMDRLMYGLAAII 59
QY 61 HGAALPLMLVFGNMTDFSPANAG-ISKNTFFVITINESTNNTQHFINLEEMTVAY 119
DB 60 HGVALPLMLVFGNMTDFSPANAG-ISKNTFFVITINESTNNTQHFINLEEMTVAY 114
QY 120 YSGIGACVLAAYIOVSFWCLAGROILKIRKOFHAIKROEIGMDVDHVGELNTRITD 179
DB 115 YTGIGACVLAAYIOVSFWCLAGROILKIRKOFHAIKROEIGMDVDHVGELNTRITD 174
QY 180 DVSFKINGIGDKIGMPFHSIATFTFGYIGTRGKMLTVLILAIISPVGLSAAIWAIIIS 239
DB 175 DVSFKINGIGDKIGMPFHSIATFTFGYIGTRGKMLTVLILAIISPVGLSAAIWAIIIS 234
QY 240 SFDDKELAAKAGAAVEVLAIRTVIAFGOKKELERKYNLEEKAGIGIKATAIANT 299
DB 235 SFDDKELAAKAGAAVEVLAIRTVIAFGOKKELERKYNLEEKAGIGIKATAIANT 294
QY 300 SIGAAFLIIYASVLAIFWYGTSLVLSSEYSGOVIYTFVSYLIGAFSIGQASPSIEAFAN 359
DB 300 SIGAAFLIIYASVLAIFWYGTSLVLSSEYSGOVIYTFVSYLIGAFSIGQASPSIEAFAN 359

DB 295 SMGAFLIIYASVLAIFWYGTSLVLSSEYSGOVIYTFVSYLIGAFSIGQASPSIEAFAN 354
QY 360 ARGAAVEIFKIIDNKSIDSYSKSGHKPNKNGNLEFKVHVHSYSRKREYVITGLNKKV 419
DB 355 ARGAAVEIFKIIDNKSIDSYSKSGHKPNKNGNLEFKVHVHSYSRKREYVITGLNKKV 414
QY 420 QSGOTVALVNSCGSGSTTYVOLMQRLYPTDGMVCIQDQIRTIIVNRRLREITGVSOEP 479
DB 415 QSGOTVALVNSCGSGSTTYVOLMQRLYPTDGMVCIQDQIRTIIVNRRLREITGVSOEP 474
QY 480 VLEFATTIAENIRYGRNVMTDEIEKAVKANAYDIFMLPKKFPPTIYGERAQLSGOKO 539
DB 475 VLEFATTIAENIRYGRNVMTDEIEKAVKANAYDIFMLPKKFPPTIYGERAQLSGOKO 534
QY 540 RIIAIAIYRNKRIILDLBATSALDTESEAVVOVALDKARKGRITVIYARLSTVRNADV 599
DB 535 RIIAIAIYRNKRIILDLBATSALDTESEAVVOVALDKARKGRITVIYARLSTVRNADV 594
QY 600 IAGFDGVIIEKGNHDELMKEKGIYFKLVMTQTRGNLEENATGESKSDALEKSPD 659
DB 595 IAGFDGVIIEKGNHDELMKEKGIYFKLVMTQTRGNLEENATGESKSDALEKSPD 654
QY 660 SSSSLIKRSTFRSIAHAGQODRKLGKEDINENVPVSEFWRIKLNSTEWPYFVYVIRC 719
DB 655 SSSSLIKRSTFRSIRGPHDQDRLSTKEALDEDPPIISFWRIKLNSTEWPYFVYVIRC 714
QY 720 AIIINGLOPAFSTIFSRIGIFTRDEDEPETKRONSNFSLFVLCIGISTFTFFLOGTFE 779
DB 715 AIIINGLOPAFSTIFSRIGIFTRDEDEPETKRONSNFSLFVLCIGISTFTFFLOGTFE 774
QY 780 GKAGGILKRIKRYVFNFSMLRDVSWFDPKNTGALTIRLANDAAQVGAIGSLRAYIT 839
DB 775 GKAGGILKRIKRYVFNFSMLRDVSWFDPKNTGALTIRLANDAAQVGAIGSLRAYIT 834
QY 840 QNIANLGTGIIISLYGMOLTLILAIIVPIATAGVVEKMLSGOALDKKLEEGAGKIA 899
DB 835 QNIANLGTGIIISLYGMOLTLILAIIVPIATAGVVEKMLSGOALDKKLEEGAGKIA 894
QY 900 TEALENRIVVSLTRDEKFEYVYAOGLQVPRYNSLRKAHIFGVSPSTIQAMMYVSACF 959
DB 895 TEALENRIVVSLTRDEKFEYVYAOGLQVPRYNSLRKAHIFGVSPSTIQAMMYVSACF 954
QY 960 RFQAYIVLANEFNFQOVLLVFEISAVIGAMAQVSSFPADYAKAKVSAHVIMITEKSP 1019
DB 955 RFQAYIVLANEFNFQOVLLVFEISAVIGAMAQVSSFPADYAKAKVSAHVIMITEKSP 1014
QY 1020 IDSYSPHGLAKPNTLEGVNTFNEVFNYPTRPDIPYLOGLSTLEKKGOTLALVSGSGGKS 1079
DB 1015 IDSYSPHGLAKPNTLEGVNTFNEVFNYPTRPDIPYLOGLSTLEKKGOTLALVSGSGGKS 1074
QY 1080 TYVOLLEFRYDPLAGSVLIDGKEIKHLVOWLRAHLGIVSOEPIILFDCSIAENIAYGDS 1139
DB 1075 TYVOLLEFRYDPLAGSVLIDGKEIKHLVOWLRAHLGIVSOEPIILFDCSIAENIAYGDS 1134
QY 1140 RYVSHETLQAAKEANIHFFETLPEKYNTRVGDGKQVLSGGOKORAIARALVROPHL 1199
DB 1135 RYVSHETLQAAKEANIHFFETLPEKYNTRVGDGKQVLSGGOKORAIARALVROPHL 1194
QY 1200 LDEDETSALDSESEKVOEALDKARREGRTCIYIAHRLSTQANDLIVYONGKVEKHGTH 1259
DB 1195 LDEDETSALDSESEKVOEALDKARREGRTCIYIAHRLSTQANDLIVYONGKVEKHGTH 1254
QY 1260 QOLLAOKGIYVSNVVOAGAKR 1281
DB 1255 QOLLAOKGIYVSNVVOAGAKR 1276

RESULT 3
A34786
multidrug resistance protein 1a - mouse
C:Species: Mus musculus (house mouse)
C:Date: 13-Jul-1990 #sequence_revision 13-Jul-1990 #text_change 02-Feb-2001
C:Accession: A34786; A35671

Query Match	Best Local Similarity	Matches 115;	Conservative	87.0%;	Score 5633;	DB 2;	Length 1276;	
				87.0%;	Pred. No. 0;			
				74;	Mismatches	85;	Indels	8; Gaps
								4;
QY	1	MDPEGGRKSGASEKFWFMGKKKKKNEKEKKKPTVSTFAMFRYSNMIDRLMYLVGMAAT	60					
Db	1	MELEDDGKGRADKRFKSMGKKSKK - EKKEKKPAVSYLTMFRYGMIDRLMYLVGTLA	59					
QY	61	HGAALPLMALVFGWNTSPANAG - ISRNKTEPPVYNESINNNQHFINHLEEDTVAAY	119					
Db	60	HGVALLPLMLLFEGMDTSPASVGNVSKNST - - - - - NNSPADKRAMFAK - LEEEMTVAYY	113					
QY	120	YSGIGAVLVAAVYIOVSWFCLAGBOLIKTKOPFHAIMROETGMPVHDVGEINTRLTD	179					
Db	114	YTGIGAVLIYAVYIOVSWFCLAGBOLIKIRKQKFFHALMDEIGFVHDVHGEINTRLTD	173					
QY	180	DYSKINEGIDGKIGMFPSIATFFGFTVGFTRGKRLTVLTLAISPVGLSAAIMATIS	239					
Db	174	DYSKINEGIDGKIGMFQAMATFFGCGFTIGTRGKRLTVLTLAISPVGLSAAIMATIS	233					
QY	240	SETDKELLAYAAAGAAVEEVLAAIRTVIAFGGQKELERYNKNIIEAAGIGIKKAITANI	299					
Db	234	SETDKELHAYAAAGAAVEEVLAAIRTVIAFGGQKELERYNNLIEAARLGIKAITANI	293					
QY	300	SIGAPELLIYASALAEFWGTSLVLSSEYSIQOVLTFPSVYLIGAFSGOASPSTEAFAN	359					
Db	294	SMGAPELLIYASALAEFWGTSLVLSKEYSIQOVLTFPSVYLIGAFSGOASPSTEAFAN	353					
QY	360	ARGAAVEIFKIIDNKRPSIDYSKSGKHPNIGNLEEFKNVHESYPSRKEVKILKGLNLKY	419					
Db	354	ARGAAVEEFLIIDNKRPSIDFSKSGKHPNIOGNLEEFKNIHESYPSRKEVKVILKGLNLKY	413					
QY	420	QSGQYVALVNSGCGKSTYQVMORLYDPTDQVNCIDGQDIRTINVRHLREITGVSOEP	479					
Db	414	KSGQYVALVNSGCGKSTYQVMORLYDPLDGNVSLDQDIRTINVRHLREITGVSOEP	473					
QY	480	VLEFATTAENIRYGRENVYMDIEKAYKEANAYDFIMKLPNKKDPTLVGERGADLSGGGO	539					
Db	474	VLEFATTAENIRYGRREDVYMDIEKAYKEANAYDFIMKLPNKKDPTLVGERGADLSGGGO	533					
QY	540	RIATARALVRNPKILLIDDEATSAIDTESSEANVQVQALDKARKGRTTIVIAHRLSTVRADY	599					
Db	534	RIATARALVRNPKILLIDDEATSAIDTESSEANVQVQALDKARKGRTTIVIAHRLSTVRADY	593					

RESULT 4
DWM51
multidrug resistance protein 1 - mouse
N:Alternate names: P-glycoprotein 1
C:Species: Mus musculus (house mouse)
C:Date: 31-Dec-1990 #sequence,revision 31-Dec-1990 #extl_change 19-Jan-2001
C:Accession: A33719; A25057; I57510
R:Raymond, M.; Gros, P.
Proc. Natl. Acad. Sci. U.S.A. 86, 6488-6492, 1989
A:title: Mammalian multidrug-resistance gene: correlation of exon organization with s
A:Reference number: A33719; MUID:69367274; PMID:2570420
A:Accession: A33719
A:Molecule type: DNA
A:Residues: 1-1276 <RAY>
R:Gros, P.; Croop, J.; Housman, D.
Cell 47, 371-380, 1986
A:title: Mammalian multidrug resistance gene: complete cDNA sequence indicates strong
A:Reference number: A25057; MUID:87028229; PMID:3768958
A:Accession: A25057
A:Molecule type: mRNA
A:Residues: 1-1276 <GRO>
A:Cross-references: GB:M14757; NID:g159100; PIDN:AAA79005.1; PID:g387426
R:Raymond, M.; Gros, P.
Mol. Cell. Biol. 10, 6036-6040, 1990

Query Match	81.8%	Score 5296	DB 1	Length 1276
Best Local Similarity	80.3%	Pred. No. 1.4e-299		
Matches 1029	Conservative 130	Mismatches 115	Indels 8	Gaps 5

Qy	660	SSSLKRRSTRSRSLHAPOGCDRLCTGEDLNNVPEVSEWRIKLKLNSTEMPPEVYGIFC	719
Db	657	SSSPL-RRSIRSVHRKODDERLSMKEAVEDEDPPLVSFFRIINLNLSEMPPLVYGIC	715
Qy	720	AINGGLQAFSIIIFSRIIGITFRDEDEPETRKONSMVSYLEIVIGIISFTTFELOGTF	779
Db	716	AVINGCIGQVFAIVSRIVGVSRDDHETKRONCNLSLFFVLWGLISFTVYFFOGTF	775
Qy	780	GAGGLIKRLRYMFRSMLRODVSWEDDPKNTGALTTRLANDAAQYKAGISRLAVT	839
Db	776	GAGGLIKRLRYMFRSMLRODVSWEDDPKNTGALTTRLANDAAQYKAGISRLAVT	835
Qy	840	QNIANIGTIIISLTYGQQLTLIAIYPIAINGVYEMKLSQALKDKKKELEGACKIA	899
Db	836	QNVALLGCVILSLTYGQQLTLIAIYPIAINGVYEMKLSQALKDKKKELEGACKIA	895
Qy	900	TEALENFTVSLAREQFEYMAQSLQVPRNSLRKAHIFGVSFTQAMMPFSYACF	959
Db	896	TEALENFTVSLAREQFEYMAQSLQVPRNSLRKAHIFGVSFTQAMMPFSYACF	955
Qy	960	RFGATVNEEMNQDVLVFSALVFGAMAVGVSSFPADYAKAKSAAHVMIIEKSPL	1019
Db	956	RFGATVNEEMNQDVLVFSALVFGAMAVGVSSFPADYAKAKSAAHVMIIEKSPL	1015
Qy	1020	IDSYSFHGLKNTLEGVNTFENEVFNPTPPDIPVLOGLSLEKKGQTLAVSSSCGKS	1079
Db	1016	IDSYTGKRLKTLLEGVNTFENEVFNPTPPDIPVLOGLSLEKKGQTLAVSSSCGKS	1075
Qy	1080	TYVQLLRFDPDLGASVLDGKEIKHLNVQMLRAHLGIYSQEBILFDCSIAENIAYGDS	1139
Db	1076	TYVQLLRFDPDLGASVLDGKEIKHLNVQMLRAHLGIYSQEBILFDCSIAENIAYGDS	1135
Qy	1140	RVVSHHEIMOAAKEANIHFIETLPEKYNRVYDCKGTOISGGCKOKIATARAIVRPHL	1199
Db	1136	RAVSHHEIVRAKEANIHFIETLPEKYNRVYDCKGTOISGGCKOKIATARAIVRPHL	1195
Qy	1200	LLEATGALDTSEKVKYQELDKARREGRTCIYAHRLSTIQNADLLVVRPNQKVEHGT	1259
Db	1196	LLEATGALDTSEKVKYQELDKARREGRTCIYAHRLSTIQNADLLVVRPNQKVEHGT	1255
Qy	1260	QQLLAQKGIYFMSVVOAGAKR 1281	
Db	1256	QQLLAQKGIYFMSVVOAGAKR 1275	
RESULT 5			
JH0502			
P-glycoprotein - rat			
N.Alternate names: multidrug resistance protein mdr1b			
C.Species: Rattus norvegicus (Norway rat)			
C.Date: 31-Mar-1992 #sequence_revision 31-Mar-1992 #text_change 02-Feb-2001			
C.Accession: JH0502; S22353; S22352			
R.Silberman, J.A.; Raunio, H.; Gant, T.W.; Thorgerirsson, S.S.			
Gene 106, 229-236, 1991			
A.Title: Cloning and characterization of a member of the rat multidrug resistance			
A.Reference number: JH0502; MUID:9203081; PMID:1662220			
A.Accession: JH0502			
A.Molecule type: mRNA			
A.Residues: 1-1277 <STL>			
A.Cross-references: GB:M62425			
R.Deuchars, K.L.; Duthie, M.; Ling, V.			
Biochim. Biophys. Acta 1130, 157-165, 1992			
A.Title: Identification of distinct P-glycoprotein gene sequences in rat.			
A.Reference number: S22351; MUID:9222089; PMID:1348630			
A.Accession: S22353			
A.Molecule type: DNA			
A.Residues: 1212-1226, 'T', 1228-1277 <DEU>			
A.Cross-references: EMBL:X61104; NID:956890; PIDN:CAA43416.1; PID:g1334219			
A.Accession: S22352			
A.Status: Preliminary			
A.Molecule type: DNA			
A.Residues: 1212-1226, 'T', 1228-1270, 'SV', 1271-1277 <DE>			
A.Cross-references: EMBL:X61103; NID:956888; PIDN:CAA43415.1; PID:g1334218			

C:Comment: This protein has the nucleotide binding motifs and ATP binding active transp

C:Genetics:

A:Gene: mdrlb

C:Superfamily: multidrug resistance protein; ATP-binding cassette homology

C:Keywords: ATP; duplication; glycoprotein; membrane protein; nucleotide binding; P-loop

F:409-603/Domain: ATP-binding cassette homology <ABC1>

F:426-434/Region: nucleotide-binding motif A (P-loop)

F:550-554/Region: nucleotide-binding motif B

F:1053-1248/Domain: ATP-binding cassette homology <ABC2>

F:1070-1078/Region: nucleotide-binding motif A (P-loop)

F:1195-1199/Region: nucleotide-binding motif B

F:432/Binding site: ATP (Lys) #status predicted

F:1076/Binding site: ATP (Lys) #status predicted

Query Match 80.7%; Score 5226.5; DB 2; Length 1277;

Best Local Similarity 79.9%; Pred. No. 1.5e-295;

Matches 1028; Conservative 131; Mismatches 112; Indels 15; Gaps 9;

QY 1 MDPEGGRKSAKNEFWKMGKSKNEKKEKPPVSTFAMFRYSNWLDRLYMVGTMALII 60

1 MEPEEGLNRAKDNKSKKSKK - - EKKEKPAVGIFGMPFRADMLDKMALGTLALII 58

DB 61 HGAALPLMALVFGNMTDSANAGISRN - KTFPVIIINESITNNTOHEIN - HLEEMETTYA 117

59 HGTILPLMLVFGYMTDSETP - - SRDPASDRATNNOSEINST - HTVSDTSLDEDMATYA 114

QY 118 YYSGIGAGLYAAVYQVFWCLAGRQILIRKQFHAIMKOEIGMPVDHVGELNTRL 177

115 YTTTGAGAGLYAVYQVSLWCLAGRQILIRKQFHAIMKOEIGMPVDHVGELNTRL 174

QY 178 TDVSKINSGIDGKIMFHSIATPFTGIVGFTFGKLTIVLISVGLSAIYAKI 237

175 TDVSKINSGIDGKIMFHSIATPFTGIVGFTFGKLTIVLISVGLSAIYAKI 234

QY 238 LSSFTKELAYAKAGVAEEVLAIRTVIAGGOKKELERYNKNLEAKGIGIKAITA 297

235 LITSTNKEIDAYAKAGVAEEVLAIRTVIAGGOKKELERYNKNLEAKGIGIKAITA 294

QY 298 NISIGAPELLIVASYALAFWYGTSLVSSFSYSGVLYVFFSVLIGAFSIGASPIEF 357

295 NISIGIVLLVYASYALAFWYGTSLVSSFSYSGVLYVFFSVLIGAFSIGAFPIEF 354

QY 358 ANAGAAVEIFKTIIDNKPISIDYSKSGHKRDNKIGKLEFNKVFHSYSRKEVKILKGLNL 417

355 ANAGAAVEIFKTIIDNKPISIDYSKSGHKRDNKIGKLEFNKVFHSYSRKEVKILKGLNL 414

QY 418 KVQSGQVYALVNGSGCKSTTVOLMORLYDPTDMVCIDGODIRITNVNHLREITGVSQ 477

415 KVQSGQVYALVNGSGCKSTTVOLMORLYDPTDMVCIDGODIRITNVNHLREITGVSQ 474

QY 478 EPLVFAITIAENIRKGRNVTMDIEKAVEANAYDPIMLPKKPTDLVGERGAOLSGGO 537

475 EPLVFAITIAENIRKGRNVTMDIEKAVEANAYDPIMLPKKPTDLVGERGAOLSGGO 534

QY 538 KORIIAALVARNPKIILLDDATSDALDTESEAVVQALDKARAGRTIYIAHLSITVRNA 594

535 KORIIAALVARNPKIILLDDATSDALDTESEAVVQALDKARAGRTIYIAHLSITVRNA 594

QY 598 DVIAFGDDGVIEGKNDHLMKEGKGYFKLYTMTGRNNEIELENAEGSESSESDALEMSP 657

595 DVIAFGDDGVIEGKNDHLMKEGKGYFKLYTMTGRNNEIELENAEGSESSESDALEMSP 654

QY 658 KNSGSLIKRRSTRSHAPGODKLGTKEDLNENVPVPSFWRIKLNTSTEMPYPVVG 717

655 KNSGSLIKRRSTRSHAPGODKLGTKEDLNENVPVPSFWRIKLNTSTEMPYPVVG 713

QY 718 FCALINGGLAPAFSIIFRIIGITREDPEPTKQNSNMVSVLVGLIISFTIFLQGF 777

714 LCAVINGCIGPFAIVFSIVGSRDDHETKQNCNLFSLPLVNGMSIFVYFQGF 773

QY 778 TGGKAGELITKRLRYWFRSMRDVSWPDKNTGTALTRLANDQAQVGAIGSLRAY 837

774 TGGKAGELITKRLRYWFRSMRDVSWPDKNTGTALTRLANDQAQVGAIGSLRAY 833

QY 838 ITONIANLGTGIIIS - LIYGMQITLLALLAIPVPIIAGVEMKMLSGOALKDKKELEGA 895

834 VTQNVANVGTGIIISLVLYGWMQITLLVPIIPLVIGTIEKMLSGOALKDKKELEIS 893

QY 896 GKITEALENFRVYSLTREQEFYMAOSLOVYRNSLRKAHIFGVFSFTQAMMYFSY 955

894 GKITEALENFRVYSLTREQEFYMAOSLOVYRNSLRKAHIFGVFSFTQAMMYFSY 953

QY 956 AGCFRFGAYLVANEPFODVLYFSALVIGANAVGVSSPADYKAKVSAHVMIIE 1015

954 AACFRRGAYLVANEPFODVLYFSALVIGANAVGVSSPADYKAKVSAHVMIIE 1013

QY 1016 KSPILDSYSPGIKKNTTEGNVTENEVFNPRPDPVLOGLSLEYKKQGTALVSSG 1075

1014 KSPILDSYSPGIKKNTTEGNVTENEVFNPRPDPVLOGLSLEYKKQGTALVSSG 1073

QY 1076 CGKSTVOLLERFYDPLAGSLVLDGKEIKLNVOMLRAHIGYSQEPILPDCIAENIAY 1135

1074 CGKSTVOLLERFYDPLAGSLVLDGKEIKLNVOCVRA - LQISQEPILPDCIAENIAY 1132

QY 1136 GDNRSVSHHEEIMQAKRANHHEITLPEKYNTRVGDKGTQLSGQKORIALARALVQ 1195

1133 GDNRSVSHHEEIMQAKRANHHEITLPEKYNTRVGDKGTQLSGQKORIALARALVQ 1192

QY 1196 PHILLDEATSLDTESEKVVQVQALDKARREGRTCIYIAHLSITQNDLVYVQNGKYRE 1255

1193 PHILLDEATSLDTESEKVVQVQALDKARREGRTCIYIAHLSITQNDLVYVQNGKYRE 1252

QY 1256 HGTQOOLAKQKIGYFSMVQVQAKR 1281

1253 HGTQOOLAKQKIGYFSMV - VQVQAKR 1276

RESULT 6

DVNS1A multidrug resistance protein 1a - mouse (fragment)

N:Alternate names: P-glycoprotein 1a

C:Species: Mus musculus (house mouse)

C>Date: 31-Dec-1990 #sequence_revision 31-Dec-1990 #text-change 19-Jan-2001

C:Accession: A34175

J:Heu, S.I.H.; Lothstein, L.; Horwitz, S.B.

A:Title: Differential overexpression of three mdrl gene family members in multidrug-re

A:Reference number: A34175; MID: 89308614; PMID: 2473069

A:Accession: A34175

A:Molecule type: mRNA

A:Residues: 1-1104 <HSD>

A:Cross-references: GB:M24417; GB:J04839; NID:9200329; PIDN:AAA03243.1; PID:9200330

C:Comment: This is an integral membrane protein overproduced in multidrug-resistant c

structurally and functionally unrelated lipophilic antitumor drugs.

C:Superfamily: multidrug resistance protein; ATP-binding cassette homology

C:Keywords: ATP; duplication; glycoprotein; nucleotide binding; P-loop; transmembrane

F:234-428/Domain: ATP-binding cassette homology <ABC1>

F:251-258/Region: nucleotide-binding motif A (P-loop)

F:375-379/Region: nucleotide-binding motif B

F:877-1073/Domain: ATP-binding cassette homology <ABC2>

F:894-901/Region: nucleotide-binding motif A (P-loop)

F:1020-1024/Region: nucleotide-binding motif B

F:1257/Binding site: ATP (Lys) #status predicted

F:900/Binding site: ATP (Lys) #status predicted

Query Match 77.0%; Score 4990; DB 1; Length 1104;

Best Local Similarity 89.2%; Pred. No. 6.9e-282;

Matches 984; Conservative 61; Mismatches 58; Indels 0; Gaps 0;

QY 179 DVNSKINGIDGKIMFHSIATPFTGIVGFTFGKLTIVLISVGLSAIYAKI 238

1 DVNSKINGIDGKIMFHSIATPFTGIVGFTFGKLTIVLISVGLSAIYAKI 60

QY 238 SFTTKEILLAAKAGVAEEVLAIRTVIAGGOKKELERYNKNLEAKGIGIKAITAN 298

61 SFTTKEILLAAKAGVAEEVLAIRTVIAGGOKKELERYNKNLEAKGIGIKAITAN 120

[illegible]

174 NTRLTVDYSKINIGDKIGMFHSHATPFTGTVGTRGMKLTIVIAISPVLGSAAI 233
174 NTRLTDLIDSKISBGIGDKVGMFOAVATPAGTIVGFIHGMKLTIVIAISPVLGSAAV 233
234 WAKLISFTDKELLAYAKAGVAEEVLAIRIVIAFGGKKELERKNKLEPAKGIK 293
234 WAKLISAFSDKELAAVAKAGVAEEALGAIRIVIAFGGKKELERKLEKREIGIK 293
294 AITANISIGAFLLIYASALAFWYGSILVSSSEYSGIVLFFVFSVLGAFSGASPS 353
294 AISANISGIAFLIYASALAFWYGSILVSKETITGNAMTYFFSLIGAFSVGAAPC 353
354 IEAFNANGAAVEIEFKIDNKPISIDYSKSGHKNLEERKNVHSPSKREYKIK 413
354 IDAFNANGAAVIEFDIIDNNPKIDFSERGHKPDISKGNLEFVDVFSYPSRANKIK 413
414 GINTKVGSGVVALVNGSGCKSTTVOLMQLYPTDGMVCINGODIRITNVHLEITG 473
414 GINTKVGSGVVALVNGSGCKSTTVOLMQLYPTDGMVCINGODIRITNVHLEITG 473
474 VVSQEPVLEATTIANIRYGRNTMDIEIEKAVEANAYDFIMKLPKPFVLVGERGAOL 533
474 VVSQEPVLEATTIANIRYGRNTMDIEIEKAVEANAYDFIMKLPKPFVLVGERGAOL 533
534 SSGOKRIALARALVNRKILLDEATSLDTESEBAYVALDKARKGRTIVIAHRLST 593
534 SSGOKRIALARALVNRKILLDEATSLDTESEBAYVALDKARKGRTIVIAHRLST 593
594 VNRADYIADFQGVYKGNHDELMKEKGIYFKLTMTGRNEIENATGESSESADAL 653
594 VNRADYIADFQGVYKGNHDELMKEKGIYFKLTMTGRNEIENATGESSESADAL 653
654 EMSPKDSCSLIKRSTRSHAPOGODRKIGTKED-LNENVPYSEWRILKLNSTEMP 712
654 EMSPKDSCSLIKRSTRSHAPOGODRKIGTKED-LNENVPYSEWRILKLNSTEMP 712
652 RAMPNKMSTLF-RHSTQKLNKNSQCKSLVDFTDGEANVPYSEWFLKLNKTEMP 710
713 FVVGICAILNGLOPASFISIRIIGTRDEDETRONMSVFLVIGIISFTF 772
713 FVVGICAILNGLOPASFISIRIIGTRDEDETRONMSVFLVIGIISFTF 772
773 FLOGFTFGKAGELTFRRLRYFRSMRLRODVSFDPKNTGALTTRLANDAAYKGAIG 832
773 FLOGFTFGKAGELTFRRLRYFRSMRLRODVSFDPKNTGALTTRLANDAAYKGAIG 832
833 SRLAVITONTANTGTGIIISLIYGMQTLTLLAIVPIITAGVYEMKMLSGOALKDKEL 892
833 SRLAVITONTANTGTGIIISLIYGMQTLTLLAIVPIITAGVYEMKMLSGOALKDKEL 892
890 EAKKITEATEIENRIVYSLTORKEFSMVVEKLYGPRNSVQKAHYGITSISOAMY 949
890 EAKKITEATEIENRIVYSLTORKEFSMVVEKLYGPRNSVQKAHYGITSISOAMY 949
953 FSYAGCFRGAIVLANEFNFDVLLVFSALYFGAMAVQVSPADYKAKAVSAHYIM 1012
953 FSYAGCFRGAIVLANEFNFDVLLVFSALYFGAMAVQVSPADYKAKAVSAHYIM 1012
950 FSYAGCFRGAIVLANEFNFDVLLVFSALYFGAMAVQVSPADYKAKAVSAHYIM 1009
950 FSYAGCFRGAIVLANEFNFDVLLVFSALYFGAMAVQVSPADYKAKAVSAHYIM 1009
1013 IIEKSPILDSYSPHGLKPNLBNVTNEVENVYPTRPDIYVLOGSLVKKGQTLATVG 1072
1013 IIEKSPILDSYSPHGLKPNLBNVTNEVENVYPTRPDIYVLOGSLVKKGQTLATVG 1072
1010 LFEKOPILDSYSEGLKPDFEGNITFNEVENVYPTRPANVPYLOGSLVKKGQTLATVG 1069
1010 LFEKOPILDSYSEGLKPDFEGNITFNEVENVYPTRPANVPYLOGSLVKKGQTLATVG 1069
1073 SSGGKSTVVOILLERFYDPLAGSVLIDGKEIKHLNVQMLRAHLGIVSQEPILFDOSIAEN 1132
1073 SSGGKSTVVOILLERFYDPLAGSVLIDGKEIKHLNVQMLRAHLGIVSQEPILFDOSIAEN 1132
1070 SSGGKSTVVOILLERFYDPLAGSVLIDGKEIKHLNVQMLRAHLGIVSQEPILFDOSIAEN 1129
1070 SSGGKSTVVOILLERFYDPLAGSVLIDGKEIKHLNVQMLRAHLGIVSQEPILFDOSIAEN 1129
1133 IAYDNRVSVSHEEIMQAKANIHETLPPEKYNTRVODKTOISGGOKRIATARAL 1192
1133 IAYDNRVSVSHEEIMQAKANIHETLPPEKYNTRVODKTOISGGOKRIATARAL 1192
1130 IAYDNRVSVSHEEIMQAKANIHETLPPEKYNTRVODKTOISGGOKRIATARAL 1189
1130 IAYDNRVSVSHEEIMQAKANIHETLPPEKYNTRVODKTOISGGOKRIATARAL 1189
1193 VROPHIILDEATSLDTESEKVOEALDKAREGRTIVIAHRLSTIONADLIVFQONK 1252
1193 VROPHIILDEATSLDTESEKVOEALDKAREGRTIVIAHRLSTIONADLIVFQONK 1252
1190 IARPOIILDEATSLDTESEKVOEALDKAREGRTIVIAHRLSTIONADLIVFQONK 1249
1190 IARPOIILDEATSLDTESEKVOEALDKAREGRTIVIAHRLSTIONADLIVFQONK 1249
1253 VKEHGTMOALLAKGIFSVMSVOAGAK 1280
1253 VKEHGTMOALLAKGIFSVMSVOAGAK 1280

DB 1250 VKEHGTMOALLAKGIFSVMSVOAGAK 1277
RESULT 8
DMS2
Multi drug resistance protein 2 - mouse
N:Alternate names: P-glycoprotein MDR2
C:Species: Mus musculus (house mouse)
C:Date: 31-Dec-1990 #sequence_revision 31-Dec-1990 #text_change 19-Jan-2001
C:Accession: A30409; S70711
R:Gross, P.; Raymond, M.; Bell, J.; Housman, D.
Mol. Cell. Biol. 8, 2770-2778, 1988
A:Title: Cloning and characterization of a second member of the mouse mdr gene family
A:Reference number: A30409; MUID:88302195; PMID:3405218
A:Accession: A30409
A:Molecule type: mRNA
A:Residues: 1-1276 <HSD>
A:Cross-references: GB:J03398; NID:g199109; PIDN:AAA39516.1; PID:g387428
R:Kirschner, L.S.
Nucleic Acids Res. 24, 2829-2834, 1996
A:Title: De novo generation of simple sequence during gene amplification.
A:Reference number: S70711; MUID:96313253; PMID:8759018
A:Accession: S70711
A:Status: nucleic acid sequence not shown; translation not shown
A:Molecule type: DNA
A:Residues: 43-92 <KIR>
A:Cross-references: EMBL:U46839; NID:g1228142; PIDN:AAC52722.1; PID:g1228143
A:Note: The nucleotide sequence was submitted to the EMBL Data Library, January 1996
C:Comment: This is an integral membrane protein overproduced in multidrug-resistant C
structurally and functionally unrelated lipophilic antitumor drugs.
C:Genetics:
A:Gene: mdr2
A:Superfamily: multidrug resistance protein; ATP-binding cassette homology
C:Keywords: ATP; duplication; glycoprotein; nucleotide binding; P-loop; transmembrane
E:1-637/633-1276/Region: duplication
E:1-637/633-1276/Region: duplication
F:409-603/Domain: ATP-binding cassette homology <ABC1>
F:426-433/Region: nucleotide-binding motif A (P-loop)
F:550-554/Region: nucleotide-binding motif B
F:1049-1245/Domain: ATP-binding cassette homology <ABC2>
F:1067-1074/Region: nucleotide-binding motif A (P-loop)
F:1192-1196/Region: nucleotide-binding motif B
F:88/94/Binding site: carbonylate (Asn) (covalent) #status predicted
F:432/Binding site: ATP (Lys) #status predicted
F:1072/Binding site: ATP (Lys) #status predicted
Query Match 75.0%; Score 4856; DB 1; Length 1276;
Best local Similarity 74.7%; Pred. No. 5.2e-274;
Matches 959; Conservative 139; Mismatches 172; Indels 14; Gaps 8;
DB 1 MDPEGRKGSAAER--NPMKMGKSKK--NEKKKKPTVSFAMFRYSNMLDRILYLVGTMA 57
DB 1 MDLEARNGTARLDDDFELGISTISNGREKKKVNILGILTFKRSMDQDLFMFLGTL 60
DB 58 AIIHGALPLTMLVFNQMTDSFA-NAGISRNKTFPIVINESITNTQHFINHLEEMTTY 116
DB 61 AIAHGSPLMLVTFPEMDKFDYDNG--NPSLPVNFSLMANGR----ILLEEMTRY 113
DB 117 AYYSGIGAVLVAAYIVQVSEWCLAAGROLIKROFRIAHNRKOTIGMFDVHDELNTR 176
DB 114 AYYSGIGAVLVAAYIVQVSEWCLAAGROLIKROFRIAHNRKOTIGMFDVHDELNTR 173
DB 177 LTDDVSKINIGIDDKIGMFHSHATPFTGTVGTRGMKLTIVIAISPVLGSAIAIMAK 236
DB 174 LTDDVSKISBGIGDKVGMFOAVATPAGTIVGFIHGMKLTIVIAISPVLGSAIVAMAK 233
DB 237 ILSFTDKELLAYAKAGVAEEVLAIRIVIAFGGKKELERKNKLEPAKGIKKAIT 296
DB 234 ILSFTDKELLAYAKAGVAEEVLAIRIVIAFGGKKELERKNKLEPAKGIKKAIS 293
DB 297 AITNIGAFLLIYASALAFWYGSILVSSSEYSGIVLFFVFSVLGAFSGASPSIRA 356
DB 294 AISNIGAFLLIYASALAFWYGSILVSKETITGNAMTYFFSLIGAFSVGAAPCIDA 353

Db 772 FLOGFTGKAGELITRLRSMFAKMLRODMSPDDYKNSGALSTRLATDRAQVQATG 831
QY 833 SRLAVTQNTANTGTGIIISLIYGMQTLTLLLAIPITIAIAGVEMKMLSGQALKDKREL 892
Db 832 TRMLAIQNTANTGTGIIISLIYGMQTLTLLLSVPFIASGIVEMKMLAGNAKDKREL 891
QY 893 EGAKRIATEAIENFRVYSLTRROKFEYMAQSLQVPRYSNLKRAHIFGVFSITQAMMY 952
Db 892 EAACKRIATEAIENFRVYSLTRROKFEYMAQSLQVPRYSNLKRAHIFGVFSITQAMMY 951
QY 953 FSYAGCRFEGYLVANFNFDVLVFSALVIGAMAVGVSSFPADYAKAVSAAHVIM 1012
Db 952 FSYAGCRFEGYLVANFNFDVLVFSALVIGAMAVGVSSFPADYAKAVSAAHVIM 1011
QY 1013 IIEKSPILDSYSPHGLKPNLTGNTVFNENYVFPTRPDIPYLGISLEVKKGQTALV 1072
Db 1012 LFERQPLIDSYSGEGLPMDKEGVSITFNEVYFNPTRANPVLQGLSLEVKKGQTALV 1071
QY 1073 SSGCGKSTVVOLLERFVDPLAGSVLIDGKEIKHLNVQMLRAHLGIYSOEPILEDCSI 1132
Db 1072 SSGCGKSTVVOLLERFVDPLAGSVLIDGKEIKHLNVQMLRAHLGIYSOEPILEDCSI 1131
QY 1133 IAYGNSRVSHHEIMQAKKANIHHEITLPEKYNTRVQDKGTQLSGCGKORITATRAL 1192
Db 1132 IAYGNSRVSHHEIMQAKKANIHHEITLPEKYNTRVQDKGTQLSGCGKORITATRAL 1191
QY 1193 VQPHILLDEATSLDTESEKVOEALDKAREGRTCIIVAHRLSTIONADLIYVQNGK 1252
Db 1192 IYQPHILLDEATSLDTESEKVOEALDKAREGRTCIIVAHRLSTIONADLIYVQNGK 1251
QY 1253 VKEHGTHQOLLQAKGIYFSMVSQVQAK 1280
Db 1252 VKEHGTHQOLLQAKGIYFSMVSQVQAK 1279

RESULT 10
S41646
P-glycoprotein - rat
C:Species: Rattus norvegicus (Norway rat)
C:Date: 13-Jan-1995 #sequence_revision 13-Jan-1995 #text_change 02-Feb-2001
C:Accession: S41646; S22354
R:Brown, P.C.; Thorgerirsson, S.S.; Siliverman, J.A.
Nucleic Acids Res. 21, 3885-3891, 1993
A:Title: Cloning and regulation of the rat mdr2 gene.
A:Reference number: S41646; MUID:93376516; PMID:8105593
A:Accession: S41646
A>Status: preliminary
A:Molecule type: mRNA
A:Residues: 11278 <BRO>
A:Cross-references: EMBL:U5079; NID:g310192; PIDN:AAA02937.1; PID:g310193
R:Deuchars, K.L.; Duthe, M.; Ling, Y.
Biochim. Biophys. Acta 1130, 157-165, 1992
A:Title: Identification of distinct P-glycoprotein gene sequences in rat.
A:Reference number: S22351; MUID:92233089; PMID:1348630
A:Accession: S22354
A>Status: preliminary
A:Molecule type: DNA
A:Residues: 1211-1278 <DEU>
A:Cross-references: EMBL:X61105; NID:g56892; PIDN:CAA43417.1; PID:g1334220
C:Superfamily: multidrug resistance protein; ATP-binding cassette homology
C:Keywords: ATP; glycoprotein; nucleotide binding; P-loop
F:409-603/domain: ATP-binding cassette homology <ABC1>
F:426-433/region: nucleotide-binding motif A (P-loop)
F:1051-1247/domain: ATP-binding cassette homology <ABC2>
F:1068-1075/region: nucleotide-binding motif A (P-loop)

Query Match 74.2%; Score 4803; DB 2; Length 1278;
Best Local Similarity 73.8%; Pred. No. 6.3e-271;
Matches 948; Conservative 146; Mismatches 177; Indels 14; Gaps 8;

QY 1 MDPEGGRGSAEK--NFKMGKSKKN-EKKEKPYSTAMPFYSWMLRLVLTGMA 57
Db 1 MDLEAARNGTARLDGDELGSISQSRKKKVNLLIGPLTFRYSMDOKRMLLGTAM 60

QY 58 AIIHGAALPLMLVFQNMVDSFA--NAGISRNKTFPVIIINESTNTNTOHFINHLEEMTY 116
Db 61 AIHAGSLPLMLVFQNMVDSFA--NFSLPVNFSLMLNGR-----ILEEMRY 113
QY 117 AYYSGIGGVLAAYVAYSFMCIAAGROILKIRKOFFHAIROEIGMFDVHDELNTR 176
Db 114 AYYSGIGGVLAAYVAYSFMCIAAGROILKIRKOFFHAIROEIGMFDVHDELNTR 173
QY 177 LTTDVSKEINIGDKIGMFESHAFETFTGFTGRKMLTIVILATIPVGLSAIAIAK 236
Db 174 LTTDVSKEINIGDKIGMFESHAFETFTGFTGRKMLTIVILATIPVGLSAIAIAK 233
QY 237 ILSFTKELLAAKGAVAEEVLAAIRTYIAFGQKKELEBYNNLEBAKGIKRAIT 296
Db 234 ILSFTKELLAAKGAVAEEVLAAIRTYIAFGQKKELEBYNNLEBAKGIKRAIT 293
QY 297 ANISGAFLIIVASYALAFWYGSITVLSSEYSIGOVLYVFSVLGAISQASPTSEA 356
Db 294 ANISGAFLIIVASYALAFWYGSITVLSSEYSIGOVLYVFSVLGAISQASPTSEA 353
QY 357 PANARGAYELFKITDKPISDSYSGHKKPDNIKGNLEFKNVHFSYPSRKEVILKGLN 416
Db 354 PANARGAYELFKITDKPISDSYSGHKKPDNIKGNLEFKNVHFSYPSRKEVILKGLN 413
QY 417 LKVGSGTVALVNGSGGKSTTVOLMORLYDPDQAVCIDGODIRPLINRHIRETVYS 476
Db 414 LKVGSGTVALVNGSGGKSTTVOLMORLYDPDQAVCIDGODIRPLINRHIRETVYS 473
QY 477 QEPVLPATTIAENIRYGENYNTMEIKAYEANAIFYIKLNPKEPTLGERGAQLSG 536
Db 474 QEPVLPATTIAENIRYGENYNTMEIKAYEANAIFYIKLNPKEPTLGERGAQLSG 533
QY 537 OKORITAAALVRNPKILLDEATSAIDTESEAVVQALDKARGRITVIAHRLSTVRN 596
Db 534 OKORITAAALVRNPKILLDEATSAIDTESEAVVQALDKARGRITVIAHRLSTVRN 593
QY 597 ADVIAGFDDGVYVEKGNDELMEKEGYEKLVTMOTRGNETELNATGESKSESALEMS 656
Db 594 ADVIAGFDDGVYVEKGNDELMEKEGYEKLVTMOTRGNETELNATGESKSESALEMS 653
QY 657 PKDSSSLIKRSTPRSIAPQODKLGFK-EDLENVNPVPSFWRIKLKSTMPYFV 715
Db 654 PKDSSSLIKRSTPRSIAPQODKLGFK-EDLENVNPVPSFWRIKLKSTMPYFV 712
QY 716 GIPCALINGLOPASITPSRIITGFTDEDEPTRYKRONMNFVILVGLIISYTFPIQ 775
Db 713 GIPCALINGLOPASITPSRIITGFTDEDEPTRYKRONMNFVILVGLIISYTFPIQ 771
QY 776 GTFEGKAGELITRLRYVFRSMLRQDVSFDDPKNTGALTTRLANDAAYKAGISRL 835
Db 772 GTFEGKAGELITRLRYVFRSMLRQDVSFDDPKNTGALTTRLANDAAYKAGISRL 831
QY 836 AVTQNTANTGTGIIISLIYGMQTLTLLLAIPITIAIAGVEMKMLSGQALKDKRELGA 895
Db 832 AVTQNTANTGTGIIISLIYGMQTLTLLLSVPFIASGIVEMKMLAGNAKDKRELGA 891
QY 896 GKATTEAIENFRVYSLTRROKFEYMAQSLQVPRYSNLKRAHIFGVFSITQAMMY 955
Db 892 GKATTEAIENFRVYSLTRROKFEYMAQSLQVPRYSNLKRAHIFGVFSITQAMMY 951
QY 956 AGCRFRFSGYLVANFNFDVLVFSALVIGAMAVGVSSFPADYAKAVSAAHVIMITE 1015
Db 952 AGCRFRFSGYLVANFNFDVLVFSALVIGAMAVGVSSFPADYAKAVSAAHVIMITE 1011
QY 1016 KSPILDSYSPHGLKPNLTGNTVFNENYVFPTRPDIPYLGISLEVKKGQTALVSGG 1075
Db 1012 KSPILDSYSPHGLKPNLTGNTVFNENYVFPTRPDIPYLGISLEVKKGQTALVSGG 1071
QY 1076 CGKSTVVOLLERFVDPLAGSVLIDGKEIKHLNVQMLRAHLGIYSOEPILEDCSI 1135
Db 1072 CGKSTVVOLLERFVDPLAGSVLIDGKEIKHLNVQMLRAHLGIYSOEPILEDCSI 1131

QY 1136 GNSNVSHSEITMOAAEANIHHFETLPEKNTVGVGDKGTOLSGGOKRIATARALVQ 1195
 Db 1132 GNSRVSVQDEIVRAAKAENIHPIETLPPKYEIRVVDKDTOLSGGOKRIATARALVQ 1191
 QY 1196 PHILLDEATLDTSESEKVVQELDKARGRCTIVIAHRLSTIQNADLIYVQNGKVE 1255
 Db 1192 PRVLLDEATLDTSESEKVVQELDKARGRCTIVIAHRLSTIQNADLIYVQNGKVE 1251
 QY 1256 HGT HQOLLAOKGIYFSWVVOAGAK 1280
 Db 1252 HGT HQOLLAOKGIYFSWVVOAGAK 1276

RESULT 11

multidrug resistance protein homolog (mdr) - African clawed frog
 C:Species: Xenopus laevis (African clawed frog)
 C>Date: 27-Oct-1995 #sequence_revision 03-Nov-1995 #text_change 02-Feb-2001
 C:Accession: S55692
 R:Castillo, G.; Shen, H.J.; Horwitz, S.B.
 Blochim. Biophys. Acta 1262, 113-123, 1995
 A>Title: A homologue of the mammalian multidrug resistance gene (mdr) is functionally ex
 A:Reference number: S55692; MUID:95322451; PMID:7599185
 A:Accession: S55692
 A>Status: preliminary
 A:Molecule type: mRNA
 A:Residues: 1-1287 <CDS>
 A>Note: the sequence from Fig. 3 is inconsistent with that from Fig. 2
 C:Superfamily: multidrug resistance protein; ATP-binding cassette homology
 C:KeyWords: ATP; nucleotide binding; P-loop
 F:420-614/Domain: ATP-binding cassette homology <ABC1>
 F:437-444/Region: nucleotide-binding motif A (P-loop)
 F:1062-1258/Domain: ATP-binding cassette homology <ABC2>
 F:1079-1086/Region: nucleotide-binding motif A (P-loop)

Query Match 68.3%; Score 4422; DB 2; Length 1287;
 Best Local Similarity 67.5%; Pred. No. 8.8e-249;

Matches 870; Conservative 172; Mismatches 212; Indels 34; Gaps 11;

QY 2 DDEGGRKSAEKNFKMKKKKKKKKPTVSTFAMFRYSNMLDRLYMLVGTMAIHH 61
 Db 21 DPNRNSK--EKKGF--SKFKKKKKETKPKVGVFTFRRSYSDKMLMFGTASLAAH 76
 QY 62 GAALPLMLVFGNMTDFANAG--ISRNRKTPVLIINSITNTQHFNLNHEEETIYAY 119
 Db 77 GAALPLMLVFGEMTDFVNVNGVDYTGFTYBESMINAS-----RELGGMTIYAY 127
 QY 120 YSGIAGVLAAYIOVSFVWCLAAQRQILIRKQPFHAIMROETIGFVDHVDGELNTRLT 179
 Db 128 YSGIAGVLAAYIOVSFVWCLAAQRQILIRKQPFHAIMROETIGFVDHVDGELNTRLT 187
 QY 180 DVSKINSGIKGIGFHSIATPFTGVTGRGKLTLYTLAISPYLGISAIAIMAKIIS 239
 Db 188 DVSKINSGIKGIMLQSLTTLVLTGFTLIGKMGKLTWVGAIISPIMGISAIAIMAKIIS 247
 QY 240 SFTDKELLATKAGAVAEVLAIRTVIAFGQKKELEKRYNKLIEBAKIGIKAKITANV 299
 Db 248 AFTNKEKLAAYKAGAVAEVLSIRTVFAFGQKKEIHREKLEBAKIGIKAKITANV 307
 QY 300 SIGAFLIYASVLAFWGSLVLSSEYISGOVLVFSVLGAFSIOQASPISEAFAN 359
 Db 308 SIGAFLIYASVLAFWGSLVLSSEYISGOVLVFSVLGAFSIOQASPISEAFAN 367
 QY 360 ARGAAVEIFKIDNKPISIDYSKSGKHNDIKNLEFFKNVHSYSPRRKVKILKGLNKV 419
 Db 368 ARGAAVEIFKIDNKPISIDYSKSGKHNDIKNLEFFKNVHSYSPRRKVKILKGLNKV 427
 QY 420 OSQGTVALVNSCGGKSTYQVLMQRLYDPLDGKVCIDGDIPTINVRHLREITGVASGP 479
 Db 428 PSQGTVALVNSCGGKSTYQVLMQRLYDPLDGKVCIDGDIPTINVRHLREITGVASGP 487
 QY 480 VLEPATTAENIRYGRNVIMDEIEKAVKANAYDFIMKIPNKFDDLVEGRGQSLSGGOK 539
 Db 480 VLEPATTAENIRYGRNVIMDEIEKAVKANAYDFIMKIPNKFDDLVEGRGQSLSGGOK 539

Db 488 ILFPTTADNIRYGRNVIMDEIEKAVKANAYDFIMKIPNKFDDLVEGRGQSLSGGOK 547
 QY 540 RIATARALVRNPKILLDEATLDTSESEKVVQELDKARAGRTTIVIAHRLSTVQNAV 599
 Db 548 RIATARALVRNPKILLDEATLDTSESEKVVQELDKARAGRTTIVIAHRLSTVQNAV 607
 QY 600 IAGFDGVIYVEKGNHDELMKEKGIYFKLYVMQT----RGNEILENATGESKSEDALEM 655
 Db 608 IAGFDGVIYVEKGNHDELMKEKGIYFKLYVMQT----RGNEILENATGESKSEDALEM 660
 QY 656 SPKDSGLIKRSTRTSIAH--PQODKRLGTRK--DLNENVPVSEWRILANSTEMPY 712
 Db 661 PVTHSHNLVRRKSSRTVTKISKVPETEDKEVDEEKKKEGPPVSFFKMYKLNKPEMPY 720
 QY 713 FVVGIFCAIINGGLOPAFSIIFRIGIFLRDEDEPK--RQNSNMPSVFLVGLIISFT 771
 Db 721 FVVGIVICAMINGATQPAFAIFRILIGVRA---GVSQKRSESSMSLFLMLAGVSFT 777
 QY 772 FFIQGFPGKAGILIKRLRYVFRSMLRDVSWEDDPKNTTGALTRTLANDAQVGA 831
 Db 778 FFIQGFPGKAGILIKRLRYVFRSMLRDVSWEDDPKNTTGALTRTLANDAQVGA 837
 QY 832 GSRLAYTQIMANLGTGIIISLYGMQLTLLATVPIIATAGVEMKLSGALKDKE 891
 Db 838 GTRLLAQNVAANLGTGIIISLYGMQLTLLATVPIIATAGVEMKLSGALKDKE 897
 QY 892 LBSAGKIATAEIENFTVSLTRQKFEYMYAOSLOVYRNSLRKAHIFGVSFTQAMM 951
 Db 898 LBSAGKIATAEIENFTVSLTRQKFEYMYAOSLOVYRNSLRKAHIFGVSFTQAMM 957
 QY 952 YFSYACGR--FGAYIVANFNFQVLLVFSIYFGAAGVSSFPADYAKARVAHV 1010
 Db 958 VFLCWVSEVLAAYLVBEGLMKLDEVFLVSSAIVGAMALQOTSSFAFDYKAMISAHI 1017
 QY 1011 IMIIEKSLIDSYSPHGLKPTLSEGNVTENFVNYPRPIPIYLOGSLSEVKGQTLAL 1070
 Db 1018 FSLERVPIDISYSPQGRPKPKNSGVNFYKGVNNTPRPIPIYLOGSLSEVKGQTLAL 1077
 QY 1071 VSSSGCGSTVQVLLERYDPLAGSVLIDKEIKHLNVQWLRHAGITVSOEPIIFDCSIA 1130
 Db 1078 VSSSGCGSTVQVLLERYDPLAGSVLIDKEIKHLNVQWLRHAGITVSOEPIIFDCSIA 1137
 QY 1131 ENIAYGDSRVVSHSEITMOAAEANIHHFETLPEKNTVGVGDKGTOLSGGOKRIATAR 1190
 Db 1138 ENIAYGDSRVVSHSEITMOAAEANIHHFETLPEKNTVGVGDKGTOLSGGOKRIATAR 1197
 QY 1191 ALYRPHILLDEATLDTSESEKVVQELDKARAGRTTIVIAHRLSTIQNADLIYVQNG 1250
 Db 1198 ALYRPHILLDEATLDTSESEKVVQELDKARAGRTTIVIAHRLSTIQNADLIYVQNG 1257
 QY 1251 GKVEHGT HQOLLAOKGIYFSWVVOAG 1278
 Db 1258 GKVEHGT HQOLLAOKGIYFSWVVOAG 1285

RESULT 12

T42228

P-glycoprotein sister - rat

N:Alternate names: multidrug resistance protein

C:Species: Rattus norvegicus (Norway rat)

C>Date: 03-Dec-1999 #sequence_revision 03-Dec-1999 #text_change 19-Jul-2002

A:Accession: T42228

R:Chilids, S.D.; Yeh, R.L.; Hui, D.; Ling, V.

Submitted to the EMBL Data Library, June 1997

A:Description: Taxol resistance mediated by the liver-specific Sister gene of P-glyco

A:Reference number: 222088

A:Accession: T42228

A>Status: preliminary; translated from GB/EMBL/DBJ

A:Molecule type: mRNA

A:Residues: 1-1321 <CDS>

A:Cross-references: EMBL:AF010597; NID:g3273483; PID:g3273484; PIDN:AAC24753.1

C:Genetics:

A:Gene: spg9

submitted to the EMBL Data Library, November 1996

A:Reference number: Z19207

A:Accession: T19982

A:Status: preliminary; translated from GB/EMBL/DBJ

A:Molecule type: DNA

A:Residues: 1-1294 <MIL>

A:Cross-references: EMBL:281484; PIDN:CAB03973.1; GSPDB:GN00023; CESP:C47A10.1

A:Experimental source: clone C47A10

R:Hasham, V.

submitted to the EMBL Data Library, April 1997

A:Reference number: Z19853

A:Accession: T24202

A:Status: preliminary; translated from GB/EMBL/DBJ

A:Molecule type: DNA

A:Residues: 1-1294 <MIL>

A:Cross-references: EMBL:293782; PIDN:CAB07855.1; GSPDB:GN00023; CESP:C47A10.1

A:Experimental source: clone C47A10

C:Genetics:

A:Gene: CESP:C47A10.1

A:Map position: 5

A:Insertions: 15/2; 67/3; 123/3; 168/2; 509/1; 649/3; 826/2; 1089/3; 1158/3; 1250/1

C:Superfamily: multidrug resistance protein; ATP-binding cassette homology

Query Match 44.1%; Score 2857; DB 2; Length 1294;

Best Local Similarity 45.6%; Pred. No. 8e-158;

Matches 579; Conservative 239; Mismatches 422; Indels 30; Gaps 9;

QY 22 SKNEKEKKPPVSVFAMFRYSNMLDRLYMLVGMALIHGAALPLMLVFGMNTDSFAN 81

Db 17 SEKKEAPPPRIPRISIPOLRYTSDRLMLAVGLIVSCANGVGLPLMSITMGVNSQNFVT 76

QY 82 AGISRKRTTPVILNISTNNTQ-----HFNLHEEMTYAAYYSIGACVLVAAYIOVS 136

Db 77 LG-----TFLDNGNSTASEKAARAFESHEVIONCLKYY-----LGCIFAFGLQNS 125

QY 137 FMCWLAAGROILKIRQPFHAIIMROELGFDVHDVGEINLTETDYSKINDGIDKIGMF 196

Db 126 CMWVCSEKSNFRQPFHSHVMOELIAYDKNTSGILSNLFLPNLEKREGTGDKGLAF 185

QY 197 HSIATFFGTFYVGTGKRLTLVLAISPVLGSLAIAWAKILSSFTDKELIYAKAGAVA 256

Db 186 QMAAFPGFVFAVFPYDWLLTLIMMSLSPFMWICGLFLAKLTATNTAKELKQYAAVAGIA 245

QY 257 EFWLAITFVIAFGQKKELERYKNKLNLEAKGIGIKKAITANISTGAELIYASTALAF 316

Db 246 EEWLTSITVIAFNGQEECKRYDIALEHGKKGKIKSFLIGAGLASFPVVIYASCIAP 305

QY 317 WYGISLVLSSEYSIGOVLTFEVSYLIGAFSIGQASPSIEAFANARAAAEIKPIIDNPS 376

Db 306 WGTNFYVSGRLSEGTIVLTFEVSVMGSMALGAGAGQGPATIGALGAAMSLYEVDRIPE 365

QY 377 IDYSKSKSHRDNLIKGNLEKRNHVESPSKEKYEIKLGNLAKVSGQVVALVGNSSGCKS 436

Db 366 IDASTSQTSPKISGRIISVKNVEFYTPRADYKILKVSGLDAPQGTATLVALGSSGCKS 425

QY 437 TTYVLMORLYDPTDGMWICIDGODIRTIINVRHLEIREITGVVSQEPVLEFATYIAENINGYEN 496

Db 426 TIILLOREFYMPDAGQILIDIDPIEDFNKIKYLRQLVQVLSQEPNLFNISTEDNITVYGS 485

QY 497 VTMEIKKAYENAAYPFMKLKPKKEDTLVGERGAOLSGGOKORITAIRALVNRNKITLL 556

Db 486 VSDIEIKALKEAQAADFIKTFPEGINTLVGRGVOMSGGOKORITAIRALVNRNKITLL 545

QY 557 DEATSAIDTSEAAVYQALDKARKGRTTIVIAHRLSTVNNANVIGFDGIVYKGNHDE 616

Db 546 DEATSAIDASESISIQSALENASRGRTTIVIAHRLSTVNRADKILVMAAGVMEGTHET 605

QY 617 LMKRKGITKLYLVMTQRGAEILENATGSSKESDALKSPKSGSSILIKRSTR----- 671

Db 606 LTRKQVGHLELVAQVFA--DVDDPKPKRKEERMSRQTSQKRGSVNFKQESQVDEPK 662

QY 672 RSIHAPGGQGRKGG--TKEDLINENVPVPSFEMILLKNTSEMPYFVYGFICAITNGCIQPA 739

Db 663 GAPPAPAAEKEIKRKELEEBEAGVAKNLFKILRVARPEMIYIFPAIIAALIGQAVMPA 722
Qy 730 FSIIFSRIGITFDEDEPETKRONSMFVSLVLYLGIIISTFTEFLQGFPGKGLTKR 789
Db 723 FSLFSPQIINIFS-NPBDQMKKDHFWALFLVLAAGVSTSMFQCSLFGVAEKLIR 781
Qy 790 LRYMFERBMLRODVSFDPDKNTGALTTRLANDAAGVKAIGSRILAVITONIANLGTGI 849
Db 782 IRSKYIRANVLRODADTFPMKHSRGTITTRATAPNIKAKILYRIGSIFNALASVGGGL 841
Qy 850 IISLIYGMQLLLLAIVPIAIIAGVEMKMLSGQALKKKELEGACKITAEIENFRY 909
Db 842 GIAFYFGQMOMFLVMAIFPPMAVQCALMMKHYGGSATSDAKEMENACKTMEAIENFRY 901
Qy 910 VSLTEOKREYVYAOISLOVPRNSLRKHAHGVGSFSTQAMMYFSYAGCFRGAIVL--A 967
Db 902 QALTLQTKLYNFCSHLAPHGNISKAITGLTYGFANSIQFTTYAARFGFLFLDK 961
Qy 968 NEFMNFODVLVFSALVFGAMAVGVSSFPADYAKAKVSAHVIMIEKSPILDSYPHG 1027
Db 962 NVLMEBENLVLFALISFSFGTIGFAASYPEYIKATFPAAGLILFNMLEEBRIDGMTSG 1021
Qy 1028 LKPNLEGVNTNEVEYENTPRTDIPVLOGLIEYKKGQTLALVSGSGCKSTVOLLER 1087
Db 1022 TVPQ--LSGEVKNKVFERYPERPAVPILOGLVNHHVKEPQTLALVGPSCGKSTYISLER 1080
Qy 1088 FVDPPLAGSLVDGKEIKHLNOMLRHLGIVSEBILDFCSIAENIAYGDSRVYSHEE 1147
Db 1081 LYPLEGATVNDNRQNNRPHLRKHALVSOEPIIDFTSIRENIYVGLQPGETTHQOI 1140
Qy 1148 MQAKKANIHETLEPEKYNTRVGDGKQOLSGGOKHAIARALVYRPHILLDEATSA 1207
Db 1141 ETACASRANIKHFIDELPDYETRVGEGKQOLSGGOKRALARALIRNPKLLDEATSA 1200
Qy 1208 LDTSESKVOEALDKAREGRTCIYAHRLSTIONADLIVFPONGKVKHGHQOLLAKG 1267
Db 1201 LDTSESKVOVALDAAKRGTCTIVAHRLSTIVNAGCIVVNGGVQGTNELIAKRG 1260
Qy 1268 IYESMVSVQA 1277
Db 1261 AYFALTKOKS 1270

RESULT 15
Db87789 protein C34G6.4 [imported] - Caenorhabditis elegans
C:Species: Caenorhabditis elegans
C:Date: 10-May-2001 #sequence_revision 10-May-2001 #text_change 17-May-2002
C:Accession: D87789
R:Anonymous, The C. elegans Sequencing Consortium.
Science 282, 2012-2018, 1998
A:Title: Genome sequence of the nematode C. elegans: a platform for investigating biolog
A:Reference number: A75000; MID:99069613; PMID:9851916
A:Note: see web sites genome.wustl.edu/gsc/C_elegans/ and www.sanger.ac.uk/Projects/C_ele
A:Note: published errata appeared in Science 283, 35, 1999; Science 283, 2103, 1999; and
A:Accession: D87789
A:Status: preliminary
A:Molecule type: DNA
A:Restrictions: 1-1289 <STO>
A:Cross-references: GB:chr_1; PIDN:AA852482.1; PID:q1943844; GSPDB:GN00019; CESP:C34G6.4
A:Gene: C34G6.4
C:Superfamily: multidrug resistance protein; ATP-binding cassette homology

Query Match 43.5%; Score 2819; Db 2; Length 1289;
Best Local Similarity 44.8%; Pred. No. 1,3e-155;
Matches 585; Conservative 249; Mismatches 397; Indels 76; Gaps 15;

Qy 21 KSKRNEK-----KEKKPTVSTFAFRFYSNNMLDRMLVGTMAAIIHG 62
Db 2 KSRNEPTVWTKPLKRNHSSDSDSIDESTVLTNYGIFYYTQGVDDLTLTGTVAAVHHG 61

Qy 63 AALPLMLVTCNMFTDSFANAGISRNKTFPY-IINESTITNNTQHFHNLHEEMTYAAVYS 121
Db 62 AGFPLLAIVLGGMVTFELRA---QNSDFVGVNVNVPBGVLVPSLDEFSEVVKCIYVL 118
Qy 122 GIGAGVLAAYIQVSWFQCLAAGROILKIRKOFHAIROEIGFVDVHDELTRITLDV 181
Db 119 VLVGVMFEFTSVYQIACFESAERLVHKLRONYLKALLROOIQMFDOQVGNLTARITDOL 178
Qy 182 SKINEIGDCKIGFHFHSIAIPEFTGYGTFRGKTLITLYLAISPVGLSAIAWAKILTSF 241
Db 179 ERVREGIDCKFALLVOMFAALAGYGVFFYSWMTLVMMGFAPLVLYSGAKSKSMAR 238
Qy 242 TDKELLAAYAKAGAAVEELAIPTVIAFGQOKKELERYNNKLEBANGIGIKATITANISI 301
Db 239 TRVEDETAAYAGALAEFEFSIRTVHSLNGKRELDFNALVEGKQGTIVKCYMGIGV 298
Qy 302 GAAPLLIYASVALAFWGTSLVLS--EYSIGOVITVFVSLIAFSGOASPIEAFANA 360
Db 299 GFSNLCMYSSVALAFWYSTLINDPFEIDRGLITVFEPAVLSSGTSLGALPHLASFGTA 358
Qy 361 RGAAYEIKIIDNKPSIDYSKSGHKPDNIKNEEFKNVHFSYPSRKEVYILGLMLKYQ 420
Db 359 RGAATVLAIVLNSHKIDPYISLEGIVDNMKGDISFDDVHRPDSRKDIHLKGISLELK 418
Qy 421 SCOTVALVNSCGKSTVOLMQLYDPTDGVICODGODITIVNRLREITGVVSGQPV 480
Db 419 AGDKIALVSGSGCKSTIVNLLQREYDPTKGRVLDGLVEVHSLRQIGIVSOEYV 478
Qy 481 LRTATIAENIRGRNVNMDIEKAVKANAYDFIMKLPNKFDLVGERAOLSGGOKOR 540
Db 479 LFDGTIYENIKKGNHNAHDQVVEACKMANANDFKRLPDGIGRVGEKQVQSGGKOR 538
Qy 541 IATARALVBNPKILLDEATSAIDTSEBAAVVOVALDKARKRTTIVIAHRLSTYRNDV 600
Db 539 IATARALVBNPKILLDEATSAIDTSEBAAVVOVALDKARKRTTIVIAHRLSTYRNDV 598
Qy 601 AGFDDGVYIEKGNHDELMKEKIVFYKLV-----TWTRGNEIELENAATESSESAL 653
Db 599 FVFKAGNIVESGSHERLMKOGIFYDMTOAOVYRQOQOQADKIDF-----DTISEAH 652
Qy 654 EMSPKDGSSLLIKRRSTRSIAHAPQODRKLCTKEDL-----NENV----- 694
Db 653 HLSRKSSTRSAI---SIATSIH-----QLAEVELIDPFNGQTNODVIRSLISAFGG 702
Qy 695 ----PVSEFWRLIKLNSTMPYFVVGIFCATTNGSLQAFSTIFSRIGITFREDPETK 750
Db 703 ECKAPPTSMKRFKFNMGDVGWFIIGIFGAFIFGVTVPFALVIAEITVNSISLAD--QM 760
Qy 751 RQNSNMFSVLFLVIGTISFTFPLQGFPGKAGEILTKRLRYMVRSMLRQDVSWFDDPK 810
Db 761 QANVYFCGMEVLMGITFVYGFFTSANCLGRGSESLTMKLRFEARKNLLRODIAFYDDL 820
Qy 811 NTTGALTTRLANDAAGVKAIGSRILAVITONIANLGTGIISLIYGMQLTLLAIVPII 870
Db 821 HGCTKICTRATDAPNVR--VYFTRLPVVLASIVTTCGALGIGIFYGMQDALTLVAVNVL 879
Qy 871 AIIAGVEMKMLSGQALKKKELEGACKITAEIENFRYVSLTREOKREYVYAOISLOVY 930
Db 880 VMGVYFCMQRHFGQKHIDTOLLEBAGKVASQAVEHRIYVHSLNRQEQHFYTCERLRPF 939
Qy 931 RNSLRKAHIGVSPSTQAMMYFSYAGCFRGAIVYANEPNFODVLVFSALVGAAMV 990
Db 940 NTNKLKHAHTYGAVFAFQSLIFPMYAAAFYLGSIIFVNOQAMPORIDYRVFAFISFCGMI 999
Qy 991 GQVSSFPADYAKAKVSAHAHYIMIEKSPILDSYSPHG--KPNLLEGVNFNEVEYENP 1049
Db 1000 GNTTSFTPDVVKARLASLFFYLHEPTPIDSLSDSGYAKPT--GNISIRNVFENYPT 1057
Qy 1050 PDIPVLOGLSLEVKGQTLALVSGSGCKSTVOLLERFYDPLAGSVLIDGKEIKHLNVO 1109
Db 1058 KQTVLQSGFTLIDKAGTVALVGHSGGCKSTINGLLERFYNOQKGMIMIDGDIRNLNIS 1117
Qy 1110 WLRAHIGIVSOEPIIDFCSIAENIAYGDSRVVSHETIMQAKKANIHETLEPEKYN 1169

Db 1118 SLRQVCTVSOEPTLTDCTIGENICGTN-RNVTQELVEAMANNHFIIGLDPGYDT 1176
Qy 1170 RVGDKGTQLSGGQKORIAIARALVROPHILLDEATSAIDPTESEKVVQEALDKAREGRTC 1229
Db 1177 HVGEKGTQLSGGQKORIAIARALVRSVLLDEATSAIDPTESEKIVQEALDAKOGRTC 1236
Qy 1230 IVIAHRLSTIONADLIIVFQNGKVEKRGTHOOLLAOKGIYFSMVSVQ 1276
Db 1237 LVIAHRLSTIONSDVIAIYSEKIVKGTNDELIRKSEIYOKFCETQ 1283

Search completed: December 9, 2002, 17:09:31
Job time : 646 secs

GenCore version 5.1.3
Copyright (c) 1993 - 2002 Compugen Ltd.

OM protein - protein search, using sw model

Run on: December 9, 2002, 16:20:54 ; Search time 290 Seconds

(without alignments)
183.211 Million cell updates/sec

Title: US-09-672-725C-2

Perfect score: 64.77

Sequence: 1 MDPGGRKGSABKNFWMKG.....LAAQGIYSWVQAGAKR 1281

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 112892 seqs, 41476328 residues

Total number of hits satisfying chosen parameters: 112892

Minimum DB seq length: 0

Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : SwissProt_40.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	5861	90.5	1280	1	MDR1_HUMAN
2	5662.5	87.4	1276	1	MDR1_CRIGR
3	5633	87.0	1276	1	MDR3_MOUSE
4	5296	81.8	1276	1	MDR1_MOUSE
5	5277	81.5	1276	1	MDR2_CRIGR
6	5226.5	80.7	1277	1	MDR1_RAT
7	4905.5	75.7	1279	1	MDR3_HUMAN
8	4856	75.0	1276	1	MDR2_MOUSE
9	4818.5	74.4	1281	1	MDR3_MOUSE
10	4803	74.2	1278	1	MDR2_CRIGR
11	3300	50.9	1321	1	MDR2_RAT
12	3224.5	49.8	1321	1	AB11_HUMAN
13	3207.5	49.5	1321	1	AB11_RABIT
14	3205.5	49.5	1321	1	AB11_MOUSE
15	2739	42.3	1321	1	AB11_RAT
16	2564.5	39.6	1302	1	MDR4_CAEL
17	2494.5	38.5	1302	1	MDR4_MOUSE
18	2345	36.2	1302	1	MDR3_MOUSE
19	2277	35.2	1324	1	MDR3_MOUSE
20	2072	32.0	1280	1	MDR1_MOUSE
21	1543	23.8	1419	1	MDR1_MOUSE
22	1216.5	18.8	1290	1	MDR1_MOUSE
23	1132	17.5	1336	1	MDR1_MOUSE
24	1018.5	15.7	1323	1	MDR1_MOUSE
25	893	13.8	715	1	MDR1_MOUSE
26	892	13.8	735	1	MDR1_MOUSE
27	886.5	13.7	738	1	MDR1_MOUSE
28	835	12.9	762	1	MDR1_MOUSE
29	828	12.8	762	1	MDR1_MOUSE
30	825	12.7	762	1	MDR1_MOUSE
31	820	12.7	766	1	MDR1_MOUSE
32	815	12.6	766	1	MDR1_MOUSE
33	770	11.9	820	1	MDL2_YEAST

ALIGNMENTS

RESULT 1	ID	MDR1_HUMAN	STANDARD:	PRT:	1280 AA.
AC	P08183	MDR1_HUMAN	012755: 014812;		
DT	01-AUG-1988	(Rel. 08, Created)			
DT	01-NOV-1997	(Rel. 35, Last sequence update)			
DT	15-JUN-2002	(Rel. 41, Last annotation update)			
DE	Multidrug resistance protein 1 (P-glycoprotein 1) (CD243 antigen).				
GN	ABCB1 OR P-glyc OR MDR1				
OS	Homo sapiens (Human).				
OC	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;				
OC	Mammalia; Eutheria; Primates; Catarrhini; Hominoidea; Homo.				
OX	NCBI-TaxID=9606;				
RP	[1]				
RP	SEQUENCE FROM N.A.				
RX	MEDLINE=87028230; PubMed=2876781;				
RA	Chen C.-J., Chin J.E., Ueda K., Clark D.P., Pastan I., Gottesman M.M.,				
RA	Roninson I.B.;				
RT	"Internal duplication and homology with bacterial transport proteins				
RT	in the mdr1 (P-glycoprotein) gene from multidrug-resistant human				
RT	cells.";				
RL	Cell 47:381-389(1986).				
RN	[2]				
RP	SEQUENCE FROM N.A.				
RX	MEDLINE=90094448; PubMed=1967175;				
RA	Chen C.-J., Clark D.P., Ueda K., Pastan I., Gottesman M.M.,				
RA	Roninson I.B.;				
RT	"Genomic organization of the human multidrug resistance (MDR1) gene				
RT	and origin of P-glycoproteins.";				
RL	J. Biol. Chem. 265:506-514(1990).				
RN	[3]				
RP	SEQUENCE FROM N.A.				
RX	MEDLINE=97190336; PubMed=9038218;				
RA	Chen G., Duran G.E., Steiger K.A., Lacayo N.J., Jaffrezou J.P.,				
RA	Dumontet C., Sikic B.I.;				
RT	"Multidrug-resistant human sarcoma cells with a mutant P-glycoprotein,				
RT	altered phenotype, and resistance to cyclosporins.";				
RL	J. Biol. Chem. 272:5974-5982(1997).				
RN	[4]				
RP	SEQUENCE OF 1-234 FROM N.A.				
RL	Smith A., Beck C., Gibson A.;				
RL	submitted (Aug-1997) to the EMBL/GenBank/DBJ databases.				
RN	[5]				
RP	SEQUENCE OF 178-215 AND 800-856 FROM N.A.				
RX	MEDLINE=90290529; PubMed=1972623;				
RA	Gekeles V., Weger S., Probst H.;				
RT	"mdr1/P-glycoprotein gene segments analyzed from various human				
RT	leukemic cell lines exhibiting different multidrug resistance				
RT	profiles.";				
RL	Biochem. Biophys. Res. Commun. 169:796-802(1990).				
RN	[6]				
RP	SEQUENCE OF 1-23 FROM N.A.				
RA	Kloka N., Tsubota J., Kakehi Y., Komano T., Gottesman M.M.,				
RA	Pastan I., Ueda K.;				
RL	submitted (Jul-1991) to the EMBL/GenBank/DBJ databases.				
RN	[7]				

34	763	11.8	598	1	Y288_THEMA
35	760	11.7	695	1	MDL1_YEAST
36	759	11.7	1437	1	MDP5_HUMAN
37	753.5	11.6	1436	1	MDP5_RAT
38	745	11.5	1545	1	MDP2_HUMAN
39	743.5	11.5	1436	1	MDP5_MOUSE
40	739.5	11.4	726	1	YFX9_SCHPO
41	734	11.3	1564	1	MDP2_RABIT
42	733	11.3	1325	1	MDP4_HUMAN
43	722	11.1	685	1	MDL1_CANAL
44	718.5	11.1	703	1	TAP2_RAT
45	708	10.9	584	1	LMRA_LUACIA

O9y44 thermotoga
P33310 saccharomyc
O15440 homo sapien
O9ymd0 rattus norv
O92887 homo sapien
O91x5 mus musculu
O28689 oryctolagus
O15439 homo sapien
P97998 candida alb
P36372 rattus norv
O9chl8 lactococcus

RP VARIANTS SER-893 AND THR-893.
RX MEDLINE-21686803; PubMed-11829140;
RA Saïto S., Iida A., Sekine A., Mura Y., Ogawa C., Kawachi S.,
Higuchi S., Nakamura Y.;
RT "Three hundred twenty-six genetic variations in gene encoding nine
members of ATP-binding cassette, subfamily B (ABCB/MDR/TAP), in the
Japanese population.";
RT Hum. Genet. 47:38-50(2002).
RL
CC -1- FUNCTION: ENERGY-DEPENDENT EFFLUX PUMP RESPONSIBLE FOR DECREASED
CC DRUG ACCUMULATION IN MULTIDRUG-RESISTANT CELLS.
CC -1- SUBCELLULAR LOCATION: Integral membrane protein.
CC -1- TISSUE SPECIFICITY: EXPRESSED IN LIVER, KIDNEY, SMALL INTESTINE
CC AND BRAIN.
CC -1- SIMILARITY: BELONGS TO THE ABC TRANSPORTER FAMILY. MDR SUBFAMILY.
CC -1- DATABASE: NAME-Atlas Genet. Cytogenet. Oncol. Haematol.;
CC WWW="http://www.infobiogen.fr/services/chromocancer/genes/PGY1D105.html".
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL; M14758; AAA59575.1; -;
DR EMBL; M29447; AAA59576.1; -;
DR EMBL; M29424; AAA59576.1; JOINED.
DR EMBL; M29425; AAA59576.1; JOINED.
DR EMBL; M29426; AAA59576.1; JOINED.
DR EMBL; M29427; AAA59576.1; JOINED.
DR EMBL; M29428; AAA59576.1; JOINED.
DR EMBL; M29429; AAA59576.1; JOINED.
DR EMBL; M29430; AAA59576.1; JOINED.
DR EMBL; M29431; AAA59576.1; JOINED.
DR EMBL; M29432; AAA59576.1; JOINED.
DR EMBL; M29433; AAA59576.1; JOINED.
DR EMBL; M29434; AAA59576.1; JOINED.
DR EMBL; M29435; AAA59576.1; JOINED.
DR EMBL; M29436; AAA59576.1; JOINED.
DR EMBL; M29437; AAA59576.1; JOINED.
DR EMBL; M29438; AAA59576.1; JOINED.
DR EMBL; M29439; AAA59576.1; JOINED.
DR EMBL; M29440; AAA59576.1; JOINED.
DR EMBL; M29441; AAA59576.1; JOINED.
DR EMBL; M29442; AAA59576.1; JOINED.
DR EMBL; M29443; AAA59576.1; JOINED.
DR EMBL; M29444; AAA59576.1; JOINED.
DR EMBL; M29445; AAA59576.1; JOINED.
DR EMBL; M29446; AAA59576.1; JOINED.
DR EMBL; AF016535; AB69423.1; -;
DR EMBL; AC002457; AAC82531.1; -;
DR EMBL; M37724; AAA88047.1; -;
DR EMBL; M37725; AAA88047.1; -;
DR EMBL; X58723; CAA15558.1; -;
DR PIR: A25059; DVHUI
DR PIR: A34914; A34914.
DR Genew; HGNC:40; ABCB1.
DR MIM; 171050; -;
DR InterPro; IPR003593; AAA_ATPase.
DR InterPro; IPR003439; ABC_transporter.
DR InterPro; IPR001140; ABCtransprtTM.
DR Pfam; PF00005; ABC_tran; 2.
DR Pfam; PF00664; ABC_membrane; 2.
DR ProDom; PD000006; ABC_transporter; 2.
DR SMART; SMO0382; AAA; 2.
DR PROSITE; PS00211; ABC_TRANSPORTER; 2.
KW ATP-binding; Glycoprotein; Transmembrane; Transport; Repeat;
KW Multigene family; Polymorphism.
KW DOMAIN 1 51 CYTOPLASMIC (POTENTIAL).
FT TRANSMEM 52 72 POTENTIAL.
FT TRANSMEM 120 140 POTENTIAL.
FT TRANSMEM 189 209 POTENTIAL.

FT	TRANSMEM	216	236	POTENTIAL.
FT	TRANSMEM	297	317	POTENTIAL.
FT	TRANSMEM	326	346	POTENTIAL.
FT	DOMAIN	347	710	CYTOPLASMIC (POTENTIAL).
FT	TRANSMEM	711	731	POTENTIAL.
FT	TRANSMEM	757	777	POTENTIAL.
FT	TRANSMEM	833	853	POTENTIAL.
FT	TRANSMEM	854	874	POTENTIAL.
FT	TRANSMEM	937	957	POTENTIAL.
FT	TRANSMEM	974	994	POTENTIAL.
FT	DOMAIN	995	1280	CYTOPLASMIC (POTENTIAL).
FT	NP_BIND	427	434	ATP (BY SIMILARITY).
FT	NP_BIND	1070	1077	ATP (BY SIMILARITY).
FT	REPEAT	1	637	
FT	REPEAT	638	1280	
FT	CARBOHYD	91	91	N-LINKED (GLCNAC. . .) (POTENTIAL).
FT	CARBOHYD	94	94	N-LINKED (GLCNAC. . .) (POTENTIAL).
FT	CARBOHYD	99	99	N-LINKED (GLCNAC. . .) (POTENTIAL).
FT	VARIANT	893	893	A -> S (IN DBSNP:2032562).
FT	VARIANT	893	893	/FTID=VAR_013361.
FT	VARIANT	893	893	A -> T.
FT	CONFLICT	23	23	/FTID=VAR_013362.
FT	CONFLICT	185	185	S -> R (IN REF. 6).
FT	CONFLICT	336	336	G -> V (IN REF. 1 AND 3).
FT	CONFLICT	412	412	MISSING (IN REF. 3).
FT	CONFLICT	438	438	G -> A (IN REF. 3).
FT	CONFLICT	438	438	Q -> S (IN REF. 3).
FT	SEQUENCE	1280 AA;	141462 MW;	ABJC279531F43675 CRC64;

Query Match 90.5%; Score 5861; DB 1; Length 1280;
Best Local Similarity 90.7%; Pred. No. 0;
Matches 1163; Conservative 55; Mismatches 60; Indels 4; Gaps 4;

QY	1	MDPESGRKGS	A-EKNFKMGKSKKNEKKKPPVSPFAMFRISNMJDRILMYGTMAI	59
DB	1	MDPESGRKGS <td>A-EKNFKMGKSKKNEKKKPPVSPFAMFRISNMJDRILMYGTMAI<td>59</td></td>	A-EKNFKMGKSKKNEKKKPPVSPFAMFRISNMJDRILMYGTMAI <td>59</td>	59
QY	60	IHGAALPLMLVFGNMTDFPANGISRNKTPVILINESITNNOHFINHLEDEMTYAY	119	
DB	60	IHGAALPLMLVFGNMTDFPANGISRNKTPVILINESITNNOHFINHLEDEMTYAY	119	
QY	120	YSGTGAAGVLAAYIOVSFWCLAAAGRIKROPHAIMROEIGMPVHYGELNRLTD	179	
DB	120	YSGTGAAGVLAAYIOVSFWCLAAAGRIKROPHAIMROEIGMPVHYGELNRLTD	179	
QY	118	YSGTGAAGVLAAYIOVSFWCLAAAGRIKROPHAIMROEIGMPVHYGELNRLTD	177	
DB	118	YSGTGAAGVLAAYIOVSFWCLAAAGRIKROPHAIMROEIGMPVHYGELNRLTD	177	
QY	180	DVSKINSGIDKIGMPHSTAFPTGCTVGTGRKGLTLVLAISPYLSIAIAKILIS	239	
DB	180	DVSKINSGIDKIGMPHSTAFPTGCTVGTGRKGLTLVLAISPYLSIAIAKILIS	239	
QY	178	DVSKINSGIDKIGMPHSTAFPTGCTVGTGRKGLTLVLAISPYLSIAIAKILIS	237	
DB	178	DVSKINSGIDKIGMPHSTAFPTGCTVGTGRKGLTLVLAISPYLSIAIAKILIS	237	
QY	240	SEPTOKELLAAYAKGAAVEEYLAIRTVAFGGOKKELERYNKNLEAKGIGIKKAITANI	299	
DB	240	SEPTOKELLAAYAKGAAVEEYLAIRTVAFGGOKKELERYNKNLEAKGIGIKKAITANI	299	
QY	300	STGAAPFLITYASYALAFWGTSLVLSSEYSIGOVLYTFEFSVLGAGSISGASPTIEFAN	359	
DB	300	STGAAPFLITYASYALAFWGTSLVLSSEYSIGOVLYTFEFSVLGAGSISGASPTIEFAN	359	
QY	298	STGAAPFLITYASYALAFWGTSLVLSSEYSIGOVLYTFEFSVLGAGSISGASPTIEFAN	357	
DB	298	STGAAPFLITYASYALAFWGTSLVLSSEYSIGOVLYTFEFSVLGAGSISGASPTIEFAN	357	
QY	360	ARGAAYELFKTIDNKRSDISYSGKGPDKNIGNERNVHSPSPSKREKLIKGLNLKY	419	
DB	360	ARGAAYELFKTIDNKRSDISYSGKGPDKNIGNERNVHSPSPSKREKLIKGLNLKY	419	
QY	420	OSGOTVALVYNSGCGKSTYVQMLORLYDPTDQVNCIDGODIRTIYNRHLEITGVVSOEP	479	
DB	420	OSGOTVALVYNSGCGKSTYVQMLORLYDPTDQVNCIDGODIRTIYNRHLEITGVVSOEP	479	
QY	418	OSGOTVALVYNSGCGKSTYVQMLORLYDPTDQVNCIDGODIRTIYNRHLEITGVVSOEP	477	
DB	418	OSGOTVALVYNSGCGKSTYVQMLORLYDPTDQVNCIDGODIRTIYNRHLEITGVVSOEP	477	
QY	480	VLFATTAENIRKGRNVIMDELEKAVYEAANAYDFIKLNKCDTLVGRGQSLGGGOKO	539	
DB	480	VLFATTAENIRKGRNVIMDELEKAVYEAANAYDFIKLNKCDTLVGRGQSLGGGOKO	539	
QY	478	VLFATTAENIRKGRNVIMDELEKAVYEAANAYDFIKLNKCDTLVGRGQSLGGGOKO	537	
DB	478	VLFATTAENIRKGRNVIMDELEKAVYEAANAYDFIKLNKCDTLVGRGQSLGGGOKO	537	
QY	540	RIATARALVNPKITLLDEATSDLTSEDAVVOVALDKARKKRTTVIAHRLSTVRADY	599	
DB	540	RIATARALVNPKITLLDEATSDLTSEDAVVOVALDKARKKRTTVIAHRLSTVRADY	599	
QY	538	RIATARALVNPKITLLDEATSDLTSEDAVVOVALDKARKKRTTVIAHRLSTVRADY	597	
DB	538	RIATARALVNPKITLLDEATSDLTSEDAVVOVALDKARKKRTTVIAHRLSTVRADY	597	

```

QY 600 IAGDDGVIVKGNHDELKKEGIYKLVMTQTRNGNELEENATGESKESDALEMSPKD 659
DB 598 IAGDDGVIVKGNHDELKKEGIYKLVMTQTRNGNELEENATGESKESDALEMSND 657
QY 660 SSSSLIKRSTRSIAPOGDRKIGTRKEDLENNPPVSWFRIKLTENMPYFVAGIYC 719
DB 658 SRSLSIKRSTRSIAPOGDRKIGTRKEDLENNPPVSWFRIKLTENMPYFVAGIYC 717
QY 720 AIIINGLOPAFSLIFSRIGITFTRDEPETKRONSNMVSFLVYIGIISFTFLQGTTF 779
DB 718 AIIINGLOPAFSLIFSRIGITFTRDEPETKRONSNMVSFLVYIGIISFTFLQGTTF 777
QY 780 GRAGELITRLKRYMFRSKMLRODVSWPDDPKNTGALTTRLANDAAQYKAGISRLAVIT 839
DB 778 GRAGELITRLKRYMFRSKMLRODVSWPDDPKNTGALTTRLANDAAQYKAGISRLAVIT 837
QY 840 ONIANIGTIIISLIYGMOLTLILAIVPITAIAGVAMKMLSQALKDKKELEGACKIA 899
DB 838 ONIANIGTIIISLIYGMOLTLILAIVPITAIAGVAMKMLSQALKDKKELEGACKIA 897
QY 900 TEAIENFTVSLTREOKFEYMYAOSLOVPRNSLRKHHIGVSEFSTIOMAMVSYAGCF 959
DB 898 TEAIENFTVSLTREOKFEYMYAOSLOVPRNSLRKHHIGVSEFSTIOMAMVSYAGCF 957
QY 960 REGAVIVANEFNPNODVILVSAIVFGAMAVGOVSFPADYAKAKVSAHVIMITEKSP 1019
DB 958 REGAVIVANEFNPNODVILVSAIVFGAMAVGOVSFPADYAKAKVSAHVIMITEKSP 1017
QY 1020 IDSYPHGKLPKPTLEGNTFNEVFNPTRPDIPLYGLSLEYKGGTALVSSSGCGKS 1079
DB 1018 IDSYPHGKLPKPTLEGNTFNEVFNPTRPDIPLYGLSLEYKGGTALVSSSGCGKS 1077
QY 1080 TVVQLEFFYPLAGSVLIDGKEIKHLNVOMLRAHLGIVSOEPILFDCSIAENIAYGDS 1139
DB 1078 TVVQLEFFYPLAGSVLIDGKEIKHLNVOMLRAHLGIVSOEPILFDCSIAENIAYGDS 1137
QY 1140 RVSHSEETMOAKENIHHEIETPEKYNRVGDGKQVLSGGCKORAIARALVROPHIL 1199
DB 1138 RVSHSEETMOAKENIHHEIETPEKYNRVGDGKQVLSGGCKORAIARALVROPHIL 1197
QY 1200 LIDETSALDTESEKVVQVQALDKAREGRTCIIVIAHRLSTIQNDLIVVONGVKEHGH 1259
DB 1198 LIDETSALDTESEKVVQVQALDKAREGRTCIIVIAHRLSTIQNDLIVVONGVKEHGH 1257
QY 1260 QOLLAQKGIYFSMYSVQAGAKR 1281
DB 1258 QOLLAQKGIYFSMYSVQAGAKR 1279

```

RESULT 2

```

MDRL_CRIGR STANDARD; PRT; 1276 AA.
AC P21448;
DT 01-NOV-1991 (Rel. 18, Created)
DT 01-NOV-1991 (Rel. 20, last sequence update)
DT 16-OCT-2001 (Rel. 40, last annotation update)
DE Multidrug resistance protein 1 (P-glycoprotein 1).
GN ABCB1 OR PGP1 OR PGP1.
OS Cricetulus griseus (Chinese hamster).
OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Cricetinae;
OC Cricetulus.
OC NCBI_Taxid=10029;
OX [1]
RN SQUENCE FROM N.A.
RX MEDLINE=92135896; Pubmed=1685679;
RT "Endicott J.A., Sarangi F., Ling V.;
RT "Complete cDNA sequences encoding the Chinese hamster P-glycoprotein
RT gene family."
RL DNA Seq. 2:89-101(1991).
RN [2]
RP SQUENCE FROM N.A.
RX MEDLINE=91154265; Pubmed=1671863;

```

```

RA Devine S.E., Hussain A., Davide J.P., Melera P.W.;
RT "Full length and alternatively spliced pgp1 transcripts in multidrug-
RT resistant Chinese hamster lung cells."
RL J. Biol. Chem. 266:4545-4555(1991).
RN [3]
RP SQUENCE OF 706-1276 FROM N.A.
RX MEDLINE=88122132; Pubmed=2893255;
RA Endicott J.A., Juranka P.F., Sarangi F., Gerlach J.H., Deuchars K.L.,
RA Ling V.;
RT "Simultaneous expression of two P-glycoprotein genes in
RT drug-sensitive Chinese hamster ovary cells."
RL Mol. Cell. Biol. 7:4075-4081(1987).
CC -I- FUNCTION: ENERGY-DEPENDENT EFFLUX PUMP RESPONSIBLE FOR DECREASED
CC DRUG ACCUMULATION IN MULTIDRUG-RESISTANT CELLS.
CC -I- SUBCELLULAR LOCATION: Integral membrane protein.
CC -I- MISCELLANEOUS: PGP ISOFORMS DIFFER IN THEIR DRUG TRANSPORT
CC CAPABILITIES: PGP1 AND PGP2 CAN MEDIATE MDR, WHILE PGP3 APPARENTLY
CC CANNOT.
CC -I- SIMILARITY: BELONGS TO THE ABC TRANSPORTER FAMILY, MDR SUBFAMILY.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation-
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL: M60040; AAA68883.1; -
DR EMBL: M59253; AAA37004.1; -
DR EMBL: M17897; AAA37006.1; -
DR PIR: A38696; DVHYIC.
DR InterPro: IPR003593; AAA_ATPase.
DR InterPro: IPR003439; ABC_transport.
DR InterPro: IPR001140; ABCtransportTM.
DR Pfam: PF00005; ABC_tran; 2.
DR Pfam: PF00664; ABC_membrane; 2.
DR ProDom: PD000006; ABC_transport; 2.
DR SMART: SM00382; ABC_transport; 2.
DR PROSITE: PS00211; ABC_TRANSPORTER; 2.
DR KAM: ATP-binding; glycoprotein; Transmembrane; Transport; Repeat;
KW Multigene family.
DR DOMAIN 1 50
FT TRANSSEM 51 71 CYTOPLASMIC (POTENTIAL).
FT TRANSSEM 117 137 POTENTIAL.
FT TRANSSEM 186 206 POTENTIAL.
FT TRANSSEM 213 233 POTENTIAL.
FT TRANSSEM 294 314 POTENTIAL.
FT TRANSSEM 323 343 POTENTIAL.
FT TRANSSEM 344 707 POTENTIAL.
FT DOMAIN 344 707 CYTOPLASMIC (POTENTIAL).
FT TRANSSEM 708 728 POTENTIAL.
FT TRANSSEM 754 774 POTENTIAL.
FT TRANSSEM 830 850 POTENTIAL.
FT TRANSSEM 851 871 POTENTIAL.
FT TRANSSEM 934 954 POTENTIAL.
FT TRANSSEM 971 991 POTENTIAL.
FT DOMAIN 992 1276 CYTOPLASMIC (POTENTIAL).
FT NP_BIND 424 431 ATP (POTENTIAL).
FT NP_BIND 1067 1074 ATP (POTENTIAL).
FT REPEAT 1 635
FT REPEAT 636 1276
FT CONFLICT 338 339
SQ SEQUENCE 1276 AA; 140925 MW; 44FF92A186B4DF CRC64;

```

Query Match 87.4%; Score 5662.5; DB 1; Length 1276;

Best Local Similarity 87.0%; Pred. No. 2e-312;

Matches 115; Conservative 77; Mismatches 83; Indels 7; Gaps 3;

```

QY 1 MDEGGRKSAKKNFMKMKKSKKEKKKPTVSTFMPFYSWMLRLVLTGVTAAIT 60
DB 1 MEFEEDFSGRKKNKFLKMKGRKSK- EKKKKPVSVFMPFYSWMLRLVLTGVTAAIT 59
QY 61 HGAALPLMMLVFGNMTDSFANAG-ISRNRKTPVLIINESTNNTQHFINHLEEMTYAAY 119

```



```

Db 60 HVALPLMLVFGMDTDSFASVGNIPINAT-----NNATQVNSDLDEGKLEEMETAYAY 114
Oy 120 YSGIGAVLVAAVYIOVSFWCLAAGRQILIKRKOFHAIIMROEIGMPYHVHVGELNRLTD 179
Db 115 YTGIGAGVLIYAVIQVSWCLAAGRQILIKRKOFHAIIMROEIGMPYHVHVGELNRLTD 174
Oy 180 DYSKINEGIGDKIGMFESIAEFGFVGTGFTGKMLTLVLAISPVLGSAAMAKILS 239
Db 175 DYSKINEGIGDKIGMFESIAEFGFVGTGFTGKMLTLVLAISPVLGSAAMAKILS 234
Oy 240 SFTDELLAYAKAGAVAEVLAIRTVIAFGOKKELEERKNKLEPAKIGIKKAITANI 239
Db 235 SFTDELLAYAKAGAVAEVLAIRTVIAFGOKKELEERKNKLEPAKIGIKKAITANI 234
Oy 300 SIGAFLLIYASALAFWGTSLVSSSEYSGOVLTFEVSILGAFSGOASPIEAFAN 359
Db 295 SMOGAFLIYASALAFWGTSLVSKESYSGOVLTFEVSILGAFSGOASPIEAFAN 354
Oy 360 ARGAAYEIKIDNKSIDSYSKSGHKPDNIKGNLEFNKVSYSRREKVLKGLMKV 419
Db 355 ARGAAYEIKIDNKSIDSYSKSGHKPDNIKGNLEFNKVSYSRREKVLKGLMKV 414
Oy 420 QSGQVALVNSGCGKSTTVOLMORLDPDGMVCIDODIRTNVRLREITGVSOEP 479
Db 415 QSGQVALVNSGCGKSTTVOLMORLDPDGMVCIDODIRTNVRLREITGVSOEP 474
Oy 480 VLEFATIAEIRYGRENVWMDLEKAVKANAYDFIMKLPNKFDVLVGERGAOLSGGOK 539
Db 475 VLEFATIAEIRYGRENVWMDLEKAVKANAYDFIMKLPNKFDVLVGERGAOLSGGOK 534
Oy 540 RIAIARALVNPRIKLLDEATSALDSEBAVVOVALDKARKRTIYIAHRLSTYRNDV 599
Db 535 RIAIARALVNPRIKLLDEATSALDSEBAVVOVALDKARKRTIYIAHRLSTYRNDV 594
Oy 600 IAGPDGCVIYKGNHDELMKEKGYFKLWMTQGRNEIELENAESESSEDALMSKPD 659
Db 595 IAGPDGCVIYKGNHDELMKEKGYFKLWMTQGRNEIELENAESESSEDALMSKPD 654
Oy 660 SSSSLIKRRSTRSINAPOGDRKLTGKEDLNEENPVPSFWKILKINSEMYEYVVGIFC 719
Db 655 SSSSLIKRRSTRSINAPOGDRKLTGKEDLNEENPVPSFWKILKINSEMYEYVVGIFC 714
Oy 720 AINGGIQAPASITFSRITIGFTRDDDEPKRQNSMFSVLVLGIIISIFTFEFGFTF 779
Db 715 AINGGIQAPASITFSRITIGFTRDDDEPKRQNSMFSVLVLGIIISIFTFEFGFTF 774
Oy 780 GKAGEILTKRLRYWFRSMLRODVSWFDPDKNTGALTTLRLANDAQAQVGAIGSLAVIT 839
Db 775 GKAGEILTKRLRYWFRSMLRODVSWFDPDKNTGALTTLRLANDAQAQVGAIGSLAVIT 834
Oy 840 QNANLGTGIIISLYQMOLTLALLAIVPIIAIAGVEMKMSGOLKXKKELEGAKIA 899
Db 835 QNANLGTGIIISLYQMOLTLALLAIVPIIAIAGVEMKMSGOLKXKKELEGAKIA 894
Oy 900 TEAENEFYVSLTREOKFEFYVNSLOVYPRNSLKAHIFGVFSITQAMTFYAGCF 959
Db 895 TEAENEFYVSLTREOKFEFYVNSLOVYPRNSLKAHIFGVFSITQAMTFYAGCF 954
Oy 960 RFGAIVLANEFMFDVLLVSAIVFGAMAVGOVSSAPYAKAKYSAHVMIIKESPL 1019
Db 955 RFGAIVLANEFMFDVLLVSAIVFGAMAVGOVSSAPYAKAKYSAHVMIIKESPL 1014
Oy 1020 IDSISPHGLKPTLEGNTEENVEVNYPTRPDIYVLOGISLEVKKGOTLALVSSGCGKS 1079
Db 1015 IDSISPHGLKPTLEGNTEENVEVNYPTRPDIYVLOGISLEVKKGOTLALVSSGCGKS 1074
Oy 1080 TVVOLLERFYDPLAGSVLIDKEIKHLNWMRLAHLGIVSOEPLDFCSIAENIAYGDN 1139
Db 1075 TVVOLLERFYDPLAGSVLIDKEIKHLNWMRLAHLGIVSOEPLDFCSIAENIAYGDN 1134
Oy 1140 RVVSHETMQAKEANIHFIETLPEKNTVSGKQVLSGGOKRIAIARALVROPHIL 1199
Db 1135 RVVSHETMQAKEANIHFIETLPEKNTVSGKQVLSGGOKRIAIARALVROPHIL 1194

```

```

Db 1135 RVVSHETMQAKEANIHFIETLPEKNTVSGKQVLSGGOKRIAIARALVROPHIL 1194
Oy 1200 LIDEATSALDSEKRVVOEALDKARKRETCIYIAHRLSTIONMDLIVONGKVKHGT 1259
Db 1195 LIDEATSALDSEKRVVOEALDKARKRETCIYIAHRLSTIONMDLIVONGKVKHGT 1254
Oy 1260 QOLLAOKGIYFSMVSVOAGAKR 1281
Db 1255 QOLLAOKGIYFSMVSVOAGAKR 1276

RESULT 3
MR3_MOUSE
ID MDR3_MOUSE STANDARD: PRT: 1276 AA.
AC P21447.1991 (Rel. 18, Created)
DT 01-MAY-1991 (Rel. 25, Last sequence update)
DT 15-JUN-2002 (Rel. 41, Last annotation update)
DE Multidrug resistance protein 3 (p-glycoprotein 3) (MDR1A).
GN ABCB1 OR ABCB4 OR PGY3 OR PGY-3 OR MDR3 OR MDR1A.
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_TaxId=10090;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=90205845; PubMed=1969610;
RA Devallet A., Gros P.;
RT "Two members of the mouse mdr gene family confer multidrug resistance
RT with overlapping but distinct drug specificities."
RL Mol. Cell. Biol. 10:1652-1663(1990).
RN [2]
RP SEQUENCE FROM N.A.
RX MEDLINE=90287150; PubMed=1972547;
RA Hsu S.I.H., Cohen D., Kirschner L.S., Lothstein L., Hartsstein M.,
RA Horwitz S.B.;
RT "Structural analysis of the mouse mdr1a (P-glycoprotein) promoter
RT reveals the basis for differential transcript heterogeneity in
RT multidrug-resistant J774.2 cells."
RL Mol. Cell. Biol. 10:3596-3606(1990).
RN [3]
RP SEQUENCE OF 173-1276 FROM N.A.
RC STRAIN=BA1B/C;
RX MEDLINE=89308614; PubMed=2473069;
RA Hsu S.I.H., Lothstein L., Horwitz S.B.;
RT "Differential overexpression of three mdr gene family members in
RT multidrug-resistant J774.2 mouse cells. Evidence that distinct P-
RT glycoprotein precursors are encoded by unique mdr genes."
RL J. Biol. Chem. 264:12053-12062(1989).
CC -1- FUNCTION: ENERGY-DEPENDENT EFFLUX PUMP RESPONSIBLE FOR DECREASED
CC DRUG ACCUMULATION IN MULTIDRUG-RESISTANT CELLS.
CC -1- SUBCELLULAR LOCATION: Integral membrane protein.
CC -1- MISCELLANEOUS: IN MOUSE THE MDR GENE FAMILY INCLUDES THREE OR MORE
CC RELATED BUT DISTINCT CELLULAR GENES.
CC -1- SIMILARITY: BELONGS TO THE ABC TRANSPORTER FAMILY. MDR SUBFAMILY.

-----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sdb.ch/announce/
CC or send an email to license@sdb-sdb.ch).
CC -----
CC EMBL: M30697; AAA39517.1; -
CC EMBL: M33581; AAA39514.1; -
CC EMBL: M33580; AAA39518.1; -
CC EMBL: M24417; AAA03243.1; -
CC PIR: A34175; DVM51A.
CC PIR: A34786; A34786.
CC MGI: MGI:97570; Abcb1a.
CC InterPro: IPR003593; AAA_ATPase.
CC InterPro: IPR003439; ABC_transporter.

```

Query Match	87.08;	Score 5633;	DB 1/	Length 1276;
Best Local Similarity	87.08;	Pred. No. 9.2e-311;		
Matches 1115;	Conservative 74;	Mismatches 85;	Indels 8;	Gaps 4
QY	1	MDPEGRGKSAEKFMFKGKSKKNEKKKPPVSTFAMFRYSNMIDRLMTVGMATIT	60	
Db	1	MELEBEDLKGADKNFMSMGKSKK-EKKEKKPAVSUTLMFRVAGWDRDRLMTVGMATIT	59	
QY	61	HGALPPLMLVFEQNMDSFNAG-ISRNNKTPVYNESITNNTOQHINHLEEMTYAY	119	
Db	60	HGALPPLMLLFEQMDTDSFASVGNVSKNST-----NMSEADRRAMPK-LDEEMTYAY	113	
QY	120	YSGIGAGVLAAAYIYVSPFCIAGROILIKROFPIAIRQELIGFVDVHDVDELNTRLTD	179	
Db	114	YTGIGAGVLVAIYIYVSPFCIAGROILIKROFPIAIRQELIGFVDVHDVDELNTRLTD	173	
QY	180	DYSKINEGIGDKIGMFFHSIAIFFETGIVGFTFGMKLTVLIAISPVGLISAATIAKLTLS	239	
Db	174	DYSKINEGIGDKIGMFFHSMATFFEGFTIIGFTFGMKLTVLIAISPVGLISAGIAKLTLS	233	
QY	240	SFTDKELIAYAKAGAAEEVLAIIRVIAFGQCKELERYKNLLEAAGIGIKKAITANI	299	
Db	234	SFTDKELHAYAKAGAAEEVLAIIRVIAFGQCKELERYNNLLEAARLIIKKAITANI	293	
QY	300	SIGAAFLLIYASIALAFWYGISLVLSSEYSIGQVLTVEFSYLIGAFSIGQASPIEAFAN	359	
Db	294	SHGAAPFLLIYASIALAFWYGTSLVISKEYSIGQVLTVEFSVLIGAFSVQASPIEAFAN	353	
QY	360	ARGAAVEYFKLIDNKPISIDSYSGSKHRDNIKGNLEFRNHFSTPSREKVKILGNLKV	419	
Db	354	ARGAAVEYFKIIDNKPISIDSFSGSKHRKDNIQGNLEFRNHFSPSKREVOILIGLMLKV	413	
QY	420	OSGQIVALVAGSGCKSTTVQLMQRLYDPTDGMVICIDGDIIRITNVHRLREITGVASQEP	479	
Db	414	KSGQIVALVAGSGCKSTTVQLMQRLYDPTDGMVISIDGDIIRITNVHRLREITGVASQEP	473	
QY	480	VLEATTIAENTRYRENTVMDIEKAYEENAYVFIMKLINKEEDTLVGERAQLSGQOK	539	

Db	474	VFPAITIAENIRYGHEDVYTMDEIEKAVEKANAYDFIMKLPHQFTLVGERGAQLSGGQKQ	533
QY	540	RIAIARALVRNPKILLDDEATSAIDPTESEAVVVALDARKKERTIVIAHRLSTVRADY	599
Db	534	RIAIARALVRNPKILLDDEATSAIDPTESEAVVVALDARKKERTIVIAHRLSTVRADY	593
QY	600	INGFDQYIVKEGNHDELMKEKGYTEKLYVTMOTRGNLELEENATGESSESDELALEMSPKD	659
Db	594	INGFDQYIVKEGNHDELMKEKGYTEKLYVTMOTRGNLELEENATGESSESDELALEMSPKD	653
QY	660	SGSSLIRKRRSTRSIIAHPQOCODRKRLGCKEDLLENYPPVSFRRILKLNSTEPPFVVGIRC	719
Db	654	SGSSLIRKRRSTRSIIAHPQOCODRKRLGCKEDLLENYPPVSFRRILKLNSTEPPFVVGIRC	713
QY	720	AIINGCLOPASTIIFSRKIGIFTRDEDPETKRQNSNMFVLEVLVIGIISFTFFLOGFTF	779
Db	714	AIINGCLOPASTIIFSRKIGIFTRDEDPETKRQNSNMFVLEVLVIGIISFTFFLOGFTF	773
QY	780	GKAGEILLTKRLRYVFRSMRLDOVSWEFDDPPKTTATLTRLANDAAQYKASRLAVIT	839
Db	774	GKAGEILLTKRLRYVFRSMRLDOVSWEFDDPPKTTATLTRLANDAAQYKASRLAVIT	833
QY	840	QININLGGIITLILYGMQLTLLILALVPIIALAGVEMKMLSGQALKDKELEGAGKIA	899
Db	834	QININLGGIITLILYGMQLTLLILALVPIIALAGVEMKMLSGQALKDKELEGAGKIA	893
QY	900	TEALENFRIVVSLTRBOKFEVYMAASLOVPYRNSLRKAHIEGVSFSTQAMMYEYACGF	959
Db	894	TEALENFRIVVSLTRBOKFEVYMAASLOVPYRNSLRKAHIEGVSFSTQAMMYEYACGF	953
QY	960	RFAGYIVLANEFNMFODVILVFSALYFGAMANGOVSSPAPYAKAKVSAHVIMTIEKSP	1019
Db	954	RFAGYIVLANEFNMFODVILVFSALYFGAMANGOVSSPAPYAKAKVSAHVIMTIEKSP	1013
QY	1020	IDISYSHGLKPTLLEGNTFENEVFNPTBPDIPLVQGLSLVEYKQOTLALVSSGGCKS	1079
Db	1014	IDISYSHGLKPTLLEGNTFENEVFNPTBPDIPLVQGLSLVEYKQOTLALVSSGGCKS	1073
QY	1080	TYVOLLERYPYDLASVYLIDKEIKELHNQWMLAHLGIVSQEPIILFDCSIAENIAYDGS	1139
Db	1074	TYVOLLERYPYDLASVYLIDKEIKELHNQWMLAHLGIVSQEPIILFDCSIAENIAYDGS	1133
QY	1140	RYSVHEELMOAKAEKANIHFETLPEKYNRVWDKGTQSGGOKRIMAIARALROPHIL	1199
Db	1134	RYSVHEELMOAKAEKANIHFETLPEKYNRVWDKGTQSGGOKRIMAIARALROPHIL	1193
QY	1200	LIDDEATSAIDPTESEAVVVALDARKKERTIVIAHRLSTIONADLIYVFONGKVEKHTH	1259
Db	1194	LIDDEATSAIDPTESEAVVVALDARKKERTIVIAHRLSTIONADLIYVFONGKVEKHTH	1253
QY	1260	QOQLAOKGIYFSMYSVOAGAKR 1281	
Db	1254	QOQLAOKGIYFSMYSVOAGAKR 1275	

RT Indicates strong homology to bacterial transport proteins.
Cell 47:371-380(1986).

RP SEQUENCE FROM N.A.
MEDLINE-89367274; PubMed-2570420;

RA Raymond M., Gros P.;
"Mammalian multidrug-resistance gene: correlation of exon
organization with structural domains and duplication of an ancestral
gene";

RT Proc. Natl. Acad. Sci. U.S.A. 86:6488-6492(1989).

RP SEQUENCE OF 1-21 FROM N.A.

RP MEDLINE-91042535; PubMed-2246681;

RA Raymond M., Gros P.;
"Cell-specific activity of cis-acting regulatory elements in the
promoter of the mouse multidrug resistance gene mdr1";

RT Mol. Cell. Biol. 10:6036-6040(1990).
-1- FUNCTION: ENERGY-DEPENDENT EFFLUX PUMP RESPONSIBLE FOR DECREASED
DRUG ACCUMULATION IN MULTIDRUG-RESISTANT CELLS.

CC -1- SUBCELLULAR LOCATION: Integral membrane protein.
CC -1- PTM: SEVERAL PHOSPHORYLATED SERINE RESIDUES ARE PRESENT IN THE
LINKER DOMAIN.

CC -1- MISCELLANEOUS: IN MOUSE THE MDR GENE FAMILY INCLUDES THREE OR MORE
RELATED BUT DISTINCT CELLULAR GENES.

CC -1- SIMILARITY: BELONGS TO THE ABC TRANSPORTER FAMILY. MDR SUBFAMILY.

CC This SWISS-PROT entry is copyright. It is produced through a collaboration
between the Swiss Institute of Bioinformatics and the EMBL outstation
at the European Bioinformatics Institute. There are no restrictions on its
use by non-profit institutions as long as its content is in no way
modified and this statement is not removed. Usage by and for commercial
entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
or send an email to license@isb-sib.ch).

CC EMBL: M14757; AAN79005.1; -

DR EMBL: M60348; AAA39513.1; -

DR PIR: A33719; DVMS1.

DR MGD: MGI:97568; Abcb1b.

DR InterPro: IPR003593; AAA_Atpase.

DR InterPro: IPR003439; ABC_transporter.

DR InterPro: IPR001140; AbctransprtTM.

DR Pfam: PF00005; ABC_tran; 2.

DR Pfam: PF00664; ABC_membrane; 2.

DR ProDom: PD000006; ABC_transportr; 2.

DR SMART: SM00382; AAA; 2.

DR PROSITE: PS00211; ABC_TRANSPORTER; 2.

KW ATP-binding; Glycoprotein; Transmembrane; Transport; Repeat;

KW Multigene family; Phosphorylation.

KW Cytoplasmic (POTENTIAL).

FT DOMAIN 1 47

FT TRANSMEM 48 71

FT TRANSMEM 119 139

FT TRANSMEM 199 219

FT TRANSMEM 296 316

FT TRANSMEM 327 347

FT DOMAIN 348 709

FT TRANSMEM 710 730

FT TRANSMEM 755 775

FT TRANSMEM 831 851

FT TRANSMEM 853 882

FT TRANSMEM 935 955

FT TRANSMEM 966 986

FT DOMAIN 987 1276

FT NP_BIND 426 433

FT NP_BIND 1068 1075

FT REPEAT 1 635

FT REPEAT 636 1276

FT CARBOHYD 73 73

FT CARBOHYD 91 91

FT CARBOHYD 96 96

FT CARBOHYD 103 103

FT SEQUENCE 1276 AA: 140993 MW: 18040DF011B0FF4E CRC64;

Query Match

81.8%; Score 5296; DB 1; Length 1276;

Best Local Similarity 80.3%; Pred. No. 1e-291;
Matches 1029; Conservative 130; Mismatches 115; Indels 8; Gaps 5;

1 MDPEGKRSKSAENKFNKMGKSKSKKNEKKRPVSTPMFYSNWDRLMLVGTMAAI 60

1 MEFEENLKRARAKNFSSKGRKSK -EKKEKPAVGVEFGMFRYADMLKLCITLTAII 59

61 HGAALPLMLVGNKNTDSANNGISRNKTPPYINESITNTQHFIN-HLEEMTYAYV 119

60 HGTLPPLMLVGNMTDSTKAEAS---ILPSITQSPNSTLIINSLEEMAIYAYV 116

120 YSGIGAVLVAAYIYVSFNCIAAGROILKIRKOFHAIKROEIGFVDVHGVNTRITD 179

117 YTGIGAGVLVAIYQVSLMCIAAGROIKIRKOFHAIKROEIGFVDVHGVNTRITD 176

180 DVSKINGEDIKGFNFHIAFFGEIYGTGRKMLTVIAISPVLGSAIAWAKILS 239

177 DVSKINDIGDKGMFFQSTTFACFIIGTSGKMLTVIAVPLGLSALWAKVLT 236

240 SFTKRELLAYAKGAAVEVLAIRTVIAPFGOKLEKRNKNEAGIGIKKAITANI 299

237 SFTKRELQAAKGAVAEVLAIKTVIAPFGOKLEKRNKNEAGIGIKKAITASI 296

300 SIGAFLIYASTALAVWGTSLVLSSEYSGOVLVFFSVLGAFTSQASPIEAFAN 359

297 SIGIAYLVVAYASVALAWYGTSLVLSSEYSGOVLVFFSVLGAFTSIGHLAPNIEAFAN 356

360 ARGAAVEFIKIIDKPSIDSKSGKPNKIGNLEKKNVHSEYSPKREKVKILKGINLKV 419

357 ARGAAVEFIKIIDKPSIDSKSGKPNKIGNLEKKNVHSEYSPKREKVKILKGINLKV 416

420 QSGQTVLVNMGSGGKSTVQVLMQRLVDPDNGVICIGODIRINRHLREITGVSOEP 479

417 KSGQTVLVNMGSGGKSTVQVLMQRLVDPDNGVICIGODIRINRHLREITGVSOEP 476

480 VFAFTTIAENIRYGRENTYMDIEKAYKANAYDFIMKLPNKFDLYLGEAGDLSGGO 539

477 VFAFTTIAENIRYGRENTYMDIEKAYKANAYDFIMKLPNKFDLYLGEAGDLSGGO 536

540 RIAIRALVNRKIIILDEATSAIDTSEANVVALDKARKGRTYIARLSTVRNADV 599

537 RIAIRALVNRKIIILDEATSAIDTSEANVVALDKARKGRTYIARLSTVRNADV 596

600 IAGFDGVIYERGNDELKREKGYEKLVTMOTRGNELTELENATGESKSESDALEMPKD 659

597 IAGFDGVIYERGNDELKREKGYEKLVTMOTRGNELTELENATGESKSESDALEMPKD 656

660 SGSSLIKRSTRSJIHAPQODRKLTGKEDLENPPVSPFRILKLNSTEMPVVGIFC 719

657 SKSPLI-RRSIYRSYVRKQDORRLSKMEKAVDEDPVLSFRILKLNSTEMPVVGIFC 715

720 AIIINGLOPAFSTIIFRIGITFRDEDPETKRONSMFVLEVLGITSFTFFLOGFTF 779

716 AVINGCLOPAFSTIIFRIGITFRDEDPETKRONSMFVLEVLGITSFTFFLOGFTF 775

780 GKAAGEILTKRLRYVRFSMLRODVSWFDPKNTGATLTRLANDAAVYKGAIGSLAIVT 839

776 GKAAGEILTKRLRYVRFSMLRODVSWFDPKNTGATLTRLANDAAVYKGAIGSLAIVT 835

840 QNANLGTGIIISLIGVMOITLLLAIVPIIAIAGVEMKMLSGALDKKLETSAGKIA 899

836 QNANLGTGIIISLIGVMOITLLLAIVPIIAIAGVEMKMLSGALDKKLETSAGKIA 895

900 TEALIEFRVTVSLTREKFEYMTAQSLOVYRNSLKRAHIFGVSPSIQAMVYFVACF 959

896 TEALIEFRVTVSLTREKFEYMTAQSLOVYRNSLKRAHIFGVSPSIQAMVYFVACF 955

960 RFAGATLVANEPANQODLVAFSAIVGAMAVGVSSFADPYAKAKVSAHVIMIEKSP 1019

956 RFAGATLVANEPANQODLVAFSAIVGAMAVGVSSFADPYAKAKVSAHVIMIEKSP 1015

1020 IDSVPGLKRNLTGKNTGVNTEVENVPTPRDIPVLGSLKPKKQTLALVSSSGCGS 1079

Db 1016 IDSYSTEGIKPTELEGNVAFNGVFNPTRPNIPIVLOGLSLEVKKGOTLALVSSGGCKS 1075
QY 1080 TVVQLERPYDPLAGSVLIDGKEIKHLVQWRAHLGIVSEPIILFDCSIENIAYGDNS 1139
Db 1076 TVVQLERPYDPLAGSVLIDGKEIKHLVQWRAHLGIVSEPIILFDCSIENIAYGDNS 1135
QY 1140 RVVSHHEIWOAKKANIHFIETLEPEKYNTRYGDKGTOLSGGOKORTAIRALVROPHIL 1199
Db 1136 RAVSHHEIWRKAKEANIHQFIDSLPDKNTRYGDKGTOLSGGOKORTAIRALVROPHIL 1195
QY 1200 ILDEATSAIDTESEKVVQVQALDKAREGTCIVIAHRLSTIQNADLIIVFONGKVKKEGTH 1259
Db 1196 ILDEATSAIDTESEKVVQVQALDKAREGTCIVIAHRLSTIQNADLIIVFONGKVKKEGTH 1255
QY 1260 QOLLAOKGIVFMSVQAGAKR 1281
Db 1256 QOLLAOKGIVFMS-VQAGAKR 1275

RESULT 5

MDR2_CRIGR STANDARD; PRT; 1276 AA.
ID MDR2_CRIGR STANDARD; PRT; 1276 AA.
AC P21449;

DT 01-MAY-1991 (Rel. 18, Created)
DT 01-NOV-1991 (Rel. 20, Last sequence update)
DT 16-OCT-2001 (Rel. 40, Last annotation update)
DE Multidrug resistance protein 2 (P-glycoprotein 2).
GN PGY2 OR PGP2.
OS Cricetus griseus (Chinese hamster).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Cricetinae;
OC Cricetus.
OX NCBI_TaxId=10029;
[1]

RP SEQUENCE FROM N.A.
RX MEDLINE=9213586; PubMed=1685679;
RA Endicott J.A., Sarangi F., Ling V.;
RT "Complete cDNA sequences encoding the Chinese hamster P-glycoprotein
RT gene family.";
RT DNA Seq. 2:89-101(1991).
RN [2]

RP SEQUENCE OF 622-1276 FROM N.A.
RX MEDLINE=88122132; PubMed=2893255;
RA Endicott J.A., Utracki P.F., Sarangi F., Gerlach J.H., Deuchars K.L.,
RA Ling V.;
RT "Simultaneous expression of two P-glycoprotein genes in
RT drug-sensitive Chinese hamster ovary cells.";
RT Mol. Cell. Biol. 7:4075-4081(1987).
CC -! FUNCTION: ENERGY-DEPENDENT EFFLUX PUMP RESPONSIBLE FOR DECREASED
CC DRUG ACCUMULATION IN MULTIDRUG-RESISTANT CELLS.
CC -! SUBCELLULAR LOCATION: Integral membrane protein.
CC -! MISCELLANEOUS: PGP ISOFORMS DIFFER IN THEIR DRUG TRANSPORT
CC CAPABILITIES: PGP1 AND PGP2 CAN MEDIATE MDR, WHILE PGP3 APPARENTLY
CC CANNOT.

CC -! SIMILARITY: BELONGS TO THE ABC TRANSPORTER FAMILY. MDR SUBFAMILY.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation-
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC or send an email to license@isb-sib.ch).

CC EMBL: M60041; AAA68884.1; -
CC EMBL: M1786; AAA37007.1; -
CC PIR: B27126; DVHY2C.
CC InterPro: IPR003593; AAA_Arpase.
CC InterPro: IPR003439; ABC_transportr.
CC InterPro: IPR001140; ABCtransportr.
CC Pfam: PF00664; ABC_tran; 2.
CC ProDom: PD000006; ABC_transportr; 2.

DR SMART: SM00382; AAA_2.
DR PROSITE: PS00211; ABC_TRANSPORTER; 2.
KW ATP-binding; Glycoprotein; Transmembrane; Transport; Repeat;
KW Multigene family.
FT DOMAIN 1 51 CYTOPLASMIC (POTENTIAL).
FT TRANSMEM 52 72 POTENTIAL.
FT TRANSMEM 119 139 POTENTIAL.
FT TRANSMEM 188 208 POTENTIAL.
FT TRANSMEM 215 235 POTENTIAL.
FT TRANSMEM 296 316 POTENTIAL.
FT TRANSMEM 325 345 POTENTIAL.
FT DOMAIN 346 708 CYTOPLASMIC (POTENTIAL).
FT TRANSMEM 709 729 POTENTIAL.
FT TRANSMEM 755 775 POTENTIAL.
FT TRANSMEM 831 851 POTENTIAL.
FT TRANSMEM 852 872 POTENTIAL.
FT TRANSMEM 935 955 POTENTIAL.
FT TRANSMEM 972 992 POTENTIAL.
FT DOMAIN 993 1276 POTENTIAL.
FT NP_BIND 426 433 CYTOPLASMIC (POTENTIAL).
FT NP_BIND 1068 1075 ATP (POTENTIAL).
FT REPEAT 1 635 ATP (POTENTIAL).
FT REPEAT 636 1276
SQ SEQUENCE 1276 AA; 141057 MW; 5096B1385628812D CRC64;

Query Match 81.5%; Score 5277; DR 1; Length 1276;
Best Local Similarity 79.9%; Pred. No. 1,2e-290;
Matches 1024; Conservative 129; Mismatches 122; Indels 6; Gaps 3;

QY 1 MDEEGKRGKGAENEFKMKKKKKKKKKPKPVSTFAFRYSNMLDRLVLTMAI 60
Db 1 MEREDPFSARADKDFLKMGRKSKKKEKKEKKNPNVIGFMRADWLDKLYVGLTAAVL 60
QY 61 HGAALPLMLVGNMTDSEFANAGISRNKTFPVINISTNTNTOHTINLLEEDMTYAY 120
Db 61 HGTSLPLMLVGNMTDSEFANAGISRNKTFPVINISTNTNTOHTINLLEEDMTYAY 117
QY 121 SGIGAGVLAATVQVSEFWCLAGROILKROFPAHROEIGFVDVDELNTRLTDD 180
Db 118 TGIAGVLAATVQVSEFWCLAGROILKROFPAHROEIGFVDVDELNTRLTDD 177
QY 181 VSKINEIGDKIGMFPFHSIATFETFTYFGTKTLTVILATISPVGLSAAIAKLSS 240
Db 178 VSKINEIGDKIGMFPFHSIATFETFTYFGTKTLTVILATISPVGLSAAIAKLSS 237
QY 241 FDKKELLVAAAGAAVEVLAIRTVIAFGQKKELERYKNLLEAKGIGKKAITANIS 300
Db 238 FDKKELLVAAAGAAVEVLAIRTVIAFGQKKELERYKNLLEAKGIGKKAITANIS 297
QY 301 IGAFLVLAAYALAFWGTSLVSESESTIGOVLTFFSVLIGAFSIGQASPIEAFANA 360
Db 298 IGAFLVLAAYALAFWGTSLVSESESTIGOVLTFFSVLIGAFSIGQASPIEAFANA 357
QY 361 RGAAYEFKTIIDNPSIDSYSKSGHKPDNTKGLDEFKNVHVSFSPKREYKILGLXQ 420
Db 358 RGAAYEFKTIIDNPSIDSYSKSGHKPDNTKGLDEFKNVHVSFSPKREYKILGLXQ 417
QY 421 SGQVLAIVGNSGCKSTTVOLMRLDPTDGMVCIIGODIRTVINHLREITGVVSEBP 480
Db 418 SGQVLAIVGNSGCKSTTVOLMRLDPTDGMVCIIGODIRTVINHLREITGVVSEBP 477
QY 481 LPAITTAENIRYGRNVTMDIEKAVKANAYDFIKLKNKFDTLVGERGAOLSGGOK 540
Db 478 LPAITTAENIRYGRNVTMDIEKAVKANAYDFIKLKNKFDTLVGERGAOLSGGOK 537
QY 541 IATAALVNRPKTLILDEATSAIDTESEAVVOYALDKARKGTTVIAHRLSTVYRNADVI 600
Db 538 IATAALVNRPKTLILDEATSAIDTESEAVVOYALDKARKGTTVIAHRLSTVYRNADVI 597
QY 601 AGDDGVIVKGNHDELMKEGIYFKLVYMQTRGNEIELENAVTGSKSESPALMSPRDS 660
Db 598 AGDDGVIVKGNHDELMKEGIYFKLVYMQTRGNEIELENAVTGSKSESPALMSPRDS 657


```

OY 418 KVSQGVAVLVNSGCGKSTTVQMLQRLDPTDGMVCIDGODIRITINVRRLREITGVSG 477
    |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
Db 415 KKSQGVAVLVNSGCGKSTTVQMLQRLDPTDGEFVSDIDODIRITINVRRLREITGVSG 474
OY 478 EPLVFATIANINIKYRGENTYMDIEKAVKAAVYDFIMKLPKPKFDTLVGERGAQLSGSG 537
    |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
Db 475 EPLVFATIANINIKYRGENTYMDIEKAVKAAVYDFIMKLPKPKFDTLVGERGAQLSGSG 534
OY 538 KORAIARALVRNPKILLDEATSAIDTESEAVVOVALDKARGRTTYIAHRLSTVRNA 597
    |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
Db 535 KORAIARALVRNPKILLDEATSAIDTESEAVVOVALDKARGRTTYIAHRLSTVRNA 594
OY 598 DVINGFDGYIVKGNHDELMKKEKGYFKLVYMQTRNGNEILEMATEGSESDALEMP 657
    |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
Db 595 DVINGFDGYIVKGNHDELMKKEKGYFKLVYMQTRNGNEILEMATEGSESDALEMP 654
OY 658 KDSGSLIKRRSTRSRTHAPOGDRKLGKEDLNENVPVSWRLIKLNSTEMPYVVG 717
    |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
Db 655 EESKSPIL-RSIRRSIRHRQDERLSKEDEVDVAVSWOIKLINISEMPYLVGV 713
OY 718 FCALINGLOPAPSIIFSRIGIFTRDEDEPTRKONSNSFVLFIYLGISITPFLOGF 777
    |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
Db 714 LCALINGLOPAPSIIFSRIGIFTRDEDEPTRKONSNSFVLFIYLGISITPFLOGF 773
OY 778 TFGKAGEILLTKRLRYWVFESMLRQVSWFDDPKNTTGALTRLANDAAQVGAIGSRVAV 837
    |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
Db 774 TFGKAGEILLTKRLRYWVFESMLRQVSWFDDPKNTTGALTRLANDAAQVGAIGSRVAV 833
OY 838 ITQNIANTGLIIS--LIYQMQLLLALAIPIAIAINGVEMKLSQALKKKELEGA 895
    |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
Db 834 VTQNVANLGTGLIISLVYGMQLLLVLIIFLVLGITIMKLLSQALKKKELEIS 893
OY 896 GKATPAIENFTVYSLRQKFEYMAQSLQVPRNSLRKAHIFGVSFSTQAMTFSY 955
    |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
Db 894 GKATPAIENFTVYSLRQKFEYMAQSLQVPRNSLRKAHIFGVSFSTQAMTFSY 953
OY 956 AACCFRGAIVLVNENFNMQDVLVSAIVFGAMAVGOVSSAPRYAAKYSAAHVIMIE 1015
    |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
Db 954 AACCFRGAIVLVNENFNMQDVLVSAIVFGAMAVGOVSSAPRYAAKYSAAHVIMIE 1013
OY 1016 KSPILDSISPHGLKPKNTLEGVNTEVENYPTRPDIPVLQGLSLLEVKKQOTLALVSSG 1075
    |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
Db 1014 KSPILDSISPHGLKPKNTLEGVNTEVENYPTRPDIPVLQGLSLLEVKKQOTLALVSSG 1073
OY 1076 CGKSTVYVLLRFRFYDPLAGSVLIDGKEIKHLNQMVLRAHGIYSOEPIFDGSAIENIAY 1135
    |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
Db 1074 CGKSTVYVLLRFRFYDPLAGSVLIDGKEIKHLNQMVLRAHGIYSOEPIFDGSAIENIAY 1132
OY 1136 GDSRVVSHHEIQAARKEANTHRIETLPEKYTRVGDGTQSGGOKORAIARALVRO 1195
    |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
Db 1133 GDSRVVSHHEIQAARKEANTHRIETLPEKYTRVGDGTQSGGOKORAIARALVRO 1192
OY 1196 PHILLDEATSAIDTESEKVOALDKARGRTTYIAHRLSTIONADLLIVFONGKVE 1255
    |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
Db 1193 PHILLDEATSAIDTESEKVOALDKARGRTTYIAHRLSTIONADLLIVFONGKVE 1252
OY 1256 HGTHQQLAOKGIYFSVAVQAGAKR 1281
    |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
Db 1253 HGTHQQLAOKGIYFSM--VQAGAKR 1276

```

```

OC Mammalia; Eutheria; Primates; Catarrhini; Hominoidea; Homo.
OX NCBI_Taxid=9606;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=89138016; PubMed=2906314;
RA van der Blik A.M., Koelman P.M., Schneider C., Borst P.;
RT "Sequence of mdr3 cDNA encoding a human P-glycoprotein.";
RL gene 71:401-411(1988).
RN [2]
RP SEQUENCE OF 856-1279 FROM N.A.
RX MEDLINE=8811519; PubMed=2892668;
RA van der Blik A.M., Baas F., ten Houte de Lange T., Koelman P.M.,
RT "The human mdr3 gene encodes a novel P-glycoprotein homologue and
RT gives rise to alternatively spliced mRNAs in liver.";
RL EMBO J. 6:3325-3331(1987).
RN [3]
RP GENE STRUCTURE.
RX MEDLINE=91161629; PubMed=2002063;
RA Lincke C.R., Smit J.J.M., van der Velde-Koerts T., Borst P.;
RT "Structure of the human MDR3 gene and physical mapping of the human
RT MDR locus.";
RL J. Biol. Chem. 266:5303-5310(1991).
CC -1- FUNCTION: ENERGY-DEPENDENT EFFLUX PUMP RESPONSIBLE FOR DECREASED
CC CAPABLE OF CONFERRING DRUG RESISTANCE. HUMAN MDR3 IS NOT
CC OF PHOSPHATIDYLCHOLINE ACROSS THE CANALICULAR MEMBRANE OF THE
CC HEPATOCYTE.
CC -1- SUBCELLULAR LOCATION: Integral membrane protein.
CC -1- DISEASE: DEFECTS IN ABCB4 ARE A CAUSE OF PROGRESSIVE FAMILIAL
CC INTRAHEPATIC CHOLESTASIS TYPE IIT (PFIC), A FORM OF ADUOSOMAL
CC RECESSIVE LIVER DISORDERS, CHARACTERIZED BY EARLY ONSET OF
CC CHOLESTASIS THAT PROGRESSES TO CIRRHOSIS AND LIVER FAILURE BEFORE
CC ADULTHOOD.
CC -1- SIMILARITY: BELONGS TO THE ABC TRANSPORTER FAMILY. MDR SUBFAMILY.
CC
CC THIS SWISS-PROT entry is copyrighted. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@sib-sib.ch).
CC
DR EMBL: M23234; AAA36207.1; -.
DR EMBL: X06181; CAA29547.1; -.
DR PIR: JS0051; DVH03.
DR PIR: A42213; A42213.
DR HSSP: P13569; INBD.
DR GENE: HGNC:45; ABCB4.
DR MIM: 171060; -.
DR MIM: 602347; -.
DR InterPro: IPR003593; AAA_ATPase.
DR InterPro: IPR003439; ABC_transporter.
DR InterPro: IPR001140; ABCtransporter.
DR Pfam: PF00005; ABC_tran; 2.
DR Pfam: PF00664; ABC_membrane; 2.
DR ProDom: PD000006; ABC_membrane; 2.
DR SMART: SM00382; AAA; 2.
DR PROSITE: PS00211; ABC_TRANSPORTER; 2.
KW ATP-binding; glycoprotein; Transmembrane; Transport; Repeat;
KW Multigene family.
FT DOMAIN 1 57 CYTOPLASMIC (POTENTIAL).
FT TRANSMEM 58 78 TRANSMEM.
FT TRANSMEM 123 143 POTENTIAL.
FT TRANSMEM 192 211 POTENTIAL.
FT TRANSMEM 216 235 POTENTIAL.
FT TRANSMEM 301 320 POTENTIAL.
FT TRANSMEM 336 354 POTENTIAL.
FT DOMAIN 355 711 CYTOPLASMIC (POTENTIAL).
FT TRANSMEM 712 732 POTENTIAL.
FT TRANSMEM 756 776 POTENTIAL.
FT TRANSMEM 832 851 POTENTIAL.

```

```

FT TRANSMEM 854 873 POTENTIAL.
FT TRANSMEM 937 956 POTENTIAL.
FT TRANSMEM 976 993 POTENTIAL.
FT TRANSMEM 994 1279 CYTOPLASMIC (POTENTIAL).
FT NP_BIND 429 436 ATP (BY SIMILARITY).
FT NP_BIND 1069 1076 ATP (BY SIMILARITY).
FT REPEAT 1 640
FT REPEAT 641 1279
FT CARBOHYD 91 91 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 97 97 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CONFLICT 1093 1093 V -> VEFDFGQ (IN REF. 2).
FT SEQUENCE 1279 AA: 140682 MW: 30589858 CD6087 CRC64;

Query Match 75.7%; Score 4905.5; DB 1; Length 1279;
Best Local Similarity 75.5%; Pred. No. 1,2e-269;
Matches 973; Conservative 131; Mismatches 165; Indels 19; Gaps 9;

QY 1 MDPEGRKG-----SAKNEWMKMKSKKNEKKEK-PTVSTFPMFRYSNMLDLMLY 53
DB 1 MDLEAKNGTAMRPSAGDE-ELGISCKOKRKKTVMKIGVLTFRYSMDQKLFMSL 59
QY 54 GTMAITIGCALPLMTLVFGMTDSFANAGISRKTEPPIINSITNTQHFINHLEEM 113
DB 60 GTIMAIAGSGPLPMIVFGEWTDKFDVT--AGNFSEPVNFSILNPKG-----ILEEN 113
QY 114 TTYAVYSGIGAVVAAYIOVSFMCIAAGROIIRKQFHAIMROEIGMDVHDYGL 173
DB 114 TTRAVYSGIGAVVAAYIOVSFMTLAGROIIRKQFHAIMROEIGMDVHDYGL 173
QY 174 NTRLTDVSKINEGIGDKIMFPHSIATFFPGFVGFGRGKLLVLAISPVLSAAL 233
DB 174 NTRLTDVSKINEGIGDKIMFPHSIATFFPGFVGFGRGKLLVLAISPVLSAAL 233
QY 174 NTRLTDVSKINEGIGDKIMFPHSIATFFPGFVGFGRGKLLVLAISPVLSAAL 233
DB 174 NTRLTDVSKINEGIGDKIMFPHSIATFFPGFVGFGRGKLLVLAISPVLSAAL 233
QY 234 VAKTIISFDEKELLVYAKAGAAVEVLAIRTVIAFGQKLEBRYKNLEAKGIGIK 293
DB 234 VAKTIISFDEKELLVYAKAGAAVEVLAIRTVIAFGQKLEBRYKNLEAKGIGIK 293
QY 294 AITANISIGAAFLIYASTALAFWTGTSIYLSSEISQIVLTPPSVILIGFSGSPS 353
DB 294 AITANISIGAAFLIYASTALAFWTGTSIYLSSEISQIVLTPPSVILIGFSGSPS 353
QY 294 AITANISIGAAFLIYASTALAFWTGTSIYLSSEISQIVLTPPSVILIGFSGSPS 353
DB 294 AITANISIGAAFLIYASTALAFWTGTSIYLSSEISQIVLTPPSVILIGFSGSPS 353
QY 354 IEAPANRGAAYIEFKIIDNKPSIDYSKSGHPNINQNLKFNKVSYSRREVKILK 413
DB 354 IEAPANRGAAYIEFKIIDNKPSIDYSKSGHPNINQNLKFNKVSYSRREVKILK 413
QY 414 GLNLIKVSQGTVALVGNSSGKSTVQMLQRLYDPLDGMVCLDGDITINVRHLREITG 473
DB 414 GLNLIKVSQGTVALVGNSSGKSTVQMLQRLYDPLDGMVCLDGDITINVRHLREITG 473
QY 474 VVSQEPVLEFNTTAENIRGRRENTVDELEKAYKEANADFTLMLPNKFDLVGERGAOL 533
DB 474 VVSQEPVLEFNTTAENIRGRRENTVDELEKAYKEANADFTLMLPNKFDLVGERGAOL 533
QY 534 SGGOKORAIARALVRNPKILLDEATISALDTESEAVVOVALDKARKGTTIVIAHRLST 593
DB 534 SGGOKORAIARALVRNPKILLDEATISALDTESEAVVOVALDKARKGTTIVIAHRLST 593
QY 594 VRNADVLAGFDGVIYVKGNDLMEKKGTYEKLVTMORGNIELENNYGESKSESADL 653
DB 594 VRNADVLAGFDGVIYVKGNDLMEKKGTYEKLVTMORGNIELENNYGESKSESADL 653
QY 594 VRNADVLAGFDGVIYVKGNDLMEKKGTYEKLVTMORGNIELENNYGESKSESADL 653
DB 594 VRNADVLAGFDGVIYVKGNDLMEKKGTYEKLVTMORGNIELENNYGESKSESADL 653
QY 654 EMSRKDSSSLIKRSTRSTIHAPOGQDRKIGTED-LINENPVPSFWRLKINSTWEPY 712
DB 654 EMSRKDSSSLIKRSTRSTIHAPOGQDRKIGTED-LINENPVPSFWRLKINSTWEPY 712
QY 713 FVVGIFCAIINGLOPAESIIFSRIGIFTRDEDEPFKRONSNMFSVLVIGIISFIF 772
DB 713 FVVGIFCAIINGLOPAESIIFSRIGIFTRDEDEPFKRONSNMFSVLVIGIISFIF 772
QY 773 FVVGIFCAIINGLOPAESIIFSRIGIFTRDEDEPFKRONSNMFSVLVIGIISFIF 772
DB 773 FVVGIFCAIINGLOPAESIIFSRIGIFTRDEDEPFKRONSNMFSVLVIGIISFIF 772

```

```

QY 833 SRLAVTQNIANIGTGITISLYGMOQLLLAIVPIIAIAGVEMKLSGOALKRREL 892
DB 830 TRLALIAONIANIGTGITISLYGMOQLLLAIVPIIAIAGVEMKLSGOALKRREL 889
QY 893 EGAGKATEAIEENFRIVSLTREOKFEYMYAOSILOVPIYNSLRKAHIGVFSFSTQAMY 952
DB 890 EAAGKATEAIEENFRIVSLTREOKFEYMYAOSILOVPIYNSLRKAHIGVFSFSTQAMY 949
QY 953 FSYAGCFRCGATVYANDEFANFODVILVFSATVFGAMAVQVSFADYAKKVSAAHYIM 1012
DB 950 FSYAGCFRCGATVYANDEFANFODVILVFSATVFGAMAVQVSFADYAKKVSAAHYIM 1009
QY 1013 IIEKSPILDSYSPHLKPNTEBGNVTFNEVENVYPTRPDIPLYLGSLSEYKKGOTLAVG 1072
DB 1010 LFEKQPLIDSYSEBGLKPKDFEGNITFNEVENVYPTRPDIPLYLGSLSEYKKGOTLAVG 1069
QY 1073 SSGGKSTVQVQLERFYPDPLAGSVLDGKEFKILANQWLRHNGIYSEPIIFDQSIEN 1132
DB 1070 SSGGKSTVQVQLERFYPDPLAGSVLDGKEFKILANQWLRHNGIYSEPIIFDQSIEN 1129
QY 1133 IAYGDSNRVYSHHEIMQAKEANIHPIETLPKRYNTRVGDGKQTLGSGOKORAIARAL 1192
DB 1130 IAYGDSNRVYSHHEIMQAKEANIHPIETLPKRYNTRVGDGKQTLGSGOKORAIARAL 1189
QY 1193 VROPHILLDEATISALDTESEKVVQELDKARREGRCITVIAHRLSTIQNADIVFQNGR 1252
DB 1190 IROPHILLDEATISALDTESEKVVQELDKARREGRCITVIAHRLSTIQNADIVFQNGR 1249
QY 1253 VKEHGHOOOLLAOKGIYFSMVSVOAGAK 1280
DB 1250 VKEHGHOOOLLAOKGIYFSMVSVOAGAK 1277

RESULT 8
MDR2_MOUSE STANDARD: PRT: 1276 AA.
ID MDR2_MOUSE
AC P21440.1991 (Rel. 18, Created)
DT 01-MAY-1991 (Rel. 18, Last sequence update)
DT 15-JUN-2002 (Rel. 41, Last annotation update)
DE Multidrug resistance protein 2 (p-glycoprotein 2).
GN ABCB4 OR PGY2 OR PGY-2 OR MDR2.
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Mus.
OC NCB1_TaxID=10090;
OX [1]
RN SEQUENCE FROM N.A.
RP MEDLINE=88302195; PubMed=3405218;
RX Gros P., Raymond M., Bell J., Housman D.;
RA "Cloning and characterization of a second member of the mouse mdr
RT gene family."
RL Mol. Cell. Biol. 8:2770-2778(1988).
RN [2]
RN SEQUENCE OF 1-23 FROM N.A.
RC SPRAIN-BALB/C;
RA Kirschner L.S., Horwitz S.B.;
RL Submitted (DEC-1991) to the EMBL/Genbank/DBJ databases.
CC -1- FUNCTION: ENERGY-DEPENDENT EFFLUX PUMP RESPONSIBLE FOR DECREASED
CC DRUG ACCUMULATION IN MULTIDRUG-RESISTANT CELLS. MOUSE MDR2 IS
CC NOT CAPABLE OF CONFERRING DRUG RESISTANCE.
CC -1- SUBCELLULAR LOCATION: Integral membrane protein.
CC -1- MISCELLANEOUS: IN MOUSE THE MDR GENE FAMILY INCLUDES THREE OR MORE
CC RELATED BUT DISTINCT CELLULAR GENES.
CC -1- SIMILARITY: BELONGS TO THE ABC TRANSPORTER FAMILY. MDR SUBFAMILY.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation
CC between the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sdb.ch/announce/
CC or send an email to license@isb-sdb.ch).

```


OY	1	MDPBGRKGSAEK--NEWMKRKSCK-NKKREKPTVTFAFMRYSNMJDLRYLVGTMA	57
		: : : : : : : : : : : : : : : : : : : : : : : :	
Dd	1	MDLEARNGTARLDJDFGLFSISNOGREGKKVNLJLTLFRYSOMOKRLFMFLGTLM	60
OY	58	AIIIGAAIPLMLVEFGMTDSFA-NAGISRNTFPVIINESITNNTOHTLNHLEEEMTY	116
		: : : : : : : : : : : : : : : : :	
Dd	61	AIANGSLPLMIVYFGEMDKFDVING---NSLPEVNESISMNPGR---ILIEEMTRY	113
OY	117	AIYYSGIAGVLVAAYIOVSFWCLAGROILIKIRQEPFHAIROMEIQMPDVHGVELLR	176
Dd	114	AIYYSGGGGVLAAYIOVSFWTLAAGRIKIRQRPFHAILREQMOPDIKOTTELNR	173
OY	177	LTDVSKINGIDCKIGMFHSIAFFMGFTVGTRGMKLTVLAISPVLGISAAITAK	236
Dd	174	LTDVSKISMSIGIDCKMGFFOALIFFGFVIGFRGMKLTVLMAISPLLGISTVMWK	233
OY	237	IILSFDTKELLAYAKAGAEEVLAIKRTVAFGOKKELERYNKNLEAKGIGIKAT	296
Dd	234	IILSFSDKEELAAAYAKAGAAEAPALRTVIAFGOKKELEBYOKHLENNAKKGIIKAIS	293
OY	297	ANISGAAPLIYASVALAFWYGTSVLVSSSEYSIGOVLYPFESVYLIGANSIGQASPSIEA	356
Dd	294	ANISMGIAFLIYASIALFMFYGSTVLYSKETTYIGNMATVFESILLGASVSQAACIDA	353
OY	357	PANAARGAAVEIFKIIDNKPSIDSYSSKSGKHNDINKMLEFKNVHSPYRKETEIKLGN	416
Dd	354	PANAARGAAVFIIDIIDNNPKIDFSERGHKPDNIKNLFEVDVHFYPBRAINIKLIKGN	413
OY	417	LKTOSOGTVALVNGSCGKSTTYOLMQRLDPTDYAMC-IDGODIRITTNRHLRETGYVS	476
Dd	414	LKTKSOGTVALVNGSCGKSTTYOLOQLDPDEGKISTIDGDIRNFNRNCIREITIGVS	473

Qy	477	QEPVLFPTTAAENIRGRGNVTMDIEIKAVKEANAYPEIMLKPNKFDVLVGERGAQLSGG	536
Db	474	QEPVLFSTTTAAENIRYGRGNVTMDIEIKAVKEANAYPEIMLKPNKFDVLVGERGAQLSGG	533
Qy	537	OKORIALAARLVANPNKILLIDDEATSLDPTESAVVQVALDARRGRTTYIAHRLSTVRN	596
Db	534	OKORIALAARLVANPNKILLIDDEATSLDPTESAEVQVALDARRGRTTYIAHRLSTVRN	593
Qy	597	ADVIAAGFDGCVIVEKGNHDELMKEKGIYFKVLTVMQTRANEITELNATNGESKESDLEMS	656
Db	594	ADVIAAGFDGCVIVEQGSHELMKKKGIVFRLVNMQTAGSQSLSEPEVELSDEKAADVA	653
Qy	657	PKHSGSSLIKRSTRSHAPQGDQRKLTGIEDLNENVPVSEWRILIKLNTSTEMPYPVYVQ	716
Db	654	P-N-GMKARIRFTRNSTKSLKSPH-QNRLEDEETNEDANVPVSEFLVKLTKTKTMEPFPVYVQ	711
Qy	717	IFCAIINGLOPAFSTIEFRIGITRDEDEETKRONSNMPSVFLVGLISTEFPFLOG	776
Db	712	TYCAIANAGALQPAFSTIISEMAITRPPDD-AVQOQKNMFLVELGLVGLSTFPFLOG	770
Qy	777	FTYFGKAGELLTKRLRYMFRSKLRQDVSWEDPKNNTTGALTTRLANDAAQVKCAIGSRLA	836
Db	771	FTYFGKAGELLTKRLRSMFAKRLRQDMSWEDPKNNTGALSTRLANDAAQVCATGTKLA	830
Qy	837	VITQNIANLGTGIIISLYQMQLTLLAIYPIIAIANGVEMMLSGOALKKLELCEAG	896
Db	831	LINQNIANLGTGIIISFYQMQLTLLSVPLTAAGIVEMMLGNNKRDKKEMEAG	890
Qy	897	KIATEAIENFTVSLTRQKFEFYVAQSLQVPRNSLRKAHIFGVSFSTIOAMMFVSA	956
Db	891	KIATEAIENIRFVSLTRQKFEFSYVEKHLGPRNSVRKAHIIYGITFISQAFMFVSA	950
Qy	957	GGCFRGATVYANEEFNRPDVLVLSAIVFGMAVGOVSPADPYAKAKVSAAHVIMILIK	1016
Db	951	GGCFRGSTVLYNGHREFADVILVSAIVGALVAGHASPADPYAKAKLSAAYILSEER	1010
Qy	1017	SPLDISYSPHGLKPMTELEGNVTFEVEVFNPTPPDIPVLOGLSLEKKKQATLAVGSGC	1076
Db	1011	QPLDISYSGEGLMPDKFEGSTFNEVEVFNPTTANPVPVLOGLSLEKKKQATLAVGSSGC	1070
Qy	1077	GKSTVYOLLERYDPLAGSVLIDGKEIKHLNVOMLRANHLGIYSQEBILFDCSIAENIYAG	1136
Db	1071	GKSTVYOLLERYDPMAGSVLIDGQEKKLNVOMLRAQIGIYSQEBILFDCSIAENIYAG	1130
Qy	1137	DNSRVVSHEEIMQAKENAHIHFEETLPKXYNTRVGDGKTOISGOKORIALAARLVANQP	1196
Db	1131	DNSRVVPHDELIVRAKAKRANHPEITLPQKYNTRVGDGKTOISGOKORIALAARLVANQP	1190
Qy	1197	HILLIDDEATSLDPTESKVVQVALDARREGRTCIYIAHRLSTTONADLIIVFONGKYEKH	1256
Db	1191	RVLIDDEATSLDPTESKVVQVALDARREGRTCIYIAHRLSTTONADLIIVTENGKYEKH	1250
Qy	1257	GTHOOLLAOKGIYFSWVSYVQAGAK	1280
Db	1251	GTHOOLLAOKGIYFSWVYIOAGIQ	1274
RESULT 9			
MDR3_CRIGR			
ID	MDR3_CRIGR	STANDARD;	PRT; 1281 AA.
AC	P23174;		
DT	01-NOV-1991 (Rel. 20, Created)		
DT	01-NOV-1991 (Rel. 20, Last sequence update)		
DT	16-Oct-2001 (Rel. 40, Last annotation update)		
DE	Multidrug resistance protein 3 (P-glycoprotein 3).		
GN	PGY3 OR PGP3.		
OS	Cricetulus griseus (Chinese hamster).		
OC	Eukaryota; Chordata; Craniata; Vertebrata; Euteleostomi;		
OC	Mammalia; Eutheria; Rodentia; Scuriognathi; Muridae; Cricetinae;		
CC	Cricetulus.		
NCBI	NCBI_TaxID=10029;		
RP	[1]		
SEQUENCE FROM N.A.			

RX MEDLINE=92135896; PubMed=1685679;
RA Endicott J A., Sarangi F., Ling V.;
RT "Complete cDNA sequences encoding the Chinese hamster P-glycoprotein
gene family.";
RL DNA Seq. 2:89-101(1991).
CC -1- FUNCTION: ENERGY-DEPENDENT EFFLUX PUMP RESPONSIBLE FOR DECREASED
CC DRUG ACCUMULATION IN MULTIDRUG-RESISTANT CELLS.
CC -1- SUBCELLULAR LOCATION: Integral membrane protein.
CC -1- MISCELLANEOUS: PGP1 AND PGP2 CAN MEDIATE MDR, WHILE PGP3 APPARENTLY
CC CAPABILITIES.
CC CANNOT.
CC -1- SIMILARITY: BELONGS TO THE ABC TRANSPORTER FAMILY. MDR SUBFAMILY.
CC
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL Outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (see <http://www.isb-sib.ch/announce/>
CC or send an email to license@isb-sib.ch).
CC
CC EMBL: M60042; AAA6885.1; -.
CC HSSP: P13569; 1NBD.
DR InterPro: IPR003593; AAA_ATPase.
DR InterPro: IPR003439; ABC_transportr.
DR InterPro: IPR001140; ABCtransprtTM.
DR Pfam: PF00005; ABC_tran; 2.
DR Pfam: PF00664; ABC_membrane; 2.
DR ProDom: PD000006; ABC_transportr; 2.
DR SMART: SM00382; AAA; 2.
DR PROSITE: PS00211; ABC_TRANSPORTER; 1.
DR ATP-binding; Glycoprotein; Transmembrane; Transport; Repeat;
KW Multigene family.
FT DOMAIN 1 57 CYTOPLASMIC (POTENTIAL).
FT TRANSSEM 58 78 POTENTIAL.
FT TRANSSEM 122 142 POTENTIAL.
FT TRANSSEM 191 211 POTENTIAL.
FT TRANSSEM 218 238 POTENTIAL.
FT TRANSSEM 299 319 POTENTIAL.
FT TRANSSEM 328 348 POTENTIAL.
FT DOMAIN 349 712 CYTOPLASMIC (POTENTIAL).
FT TRANSSEM 713 733 POTENTIAL.
FT TRANSSEM 758 778 POTENTIAL.
FT TRANSSEM 834 854 POTENTIAL.
FT TRANSSEM 855 875 POTENTIAL.
FT TRANSSEM 938 958 POTENTIAL.
FT TRANSSEM 975 995 POTENTIAL.
FT TRANSSEM 996 1281 CYTOPLASMIC (POTENTIAL).
FT NP_BIND 429 436 ATP (POTENTIAL).
FT NP_BIND 1071 1078 ATP (POTENTIAL).
SO SEQUENCE 1281 AA; 140866 MW; 2203EF61BB29602 CRC64;

Query Match 74.4%; Score 4818.5; DB 1; Length 1281;
Best Local Similarity 73.6%; Pred. No. 1e-264;
Matches 948; Conservative 147; Mismatches 176; Indels 17; Gaps 8;

OY 1 MDPEGRGGSAAK-----NFMKMGKSKKNKKEKKKPT-VSTAMPEYSMWLRLVWLVG 54
DB 1 MDLEAARNGTARPPGTVEGDFELFELISISQGNKKKYNLIGPLTLFRTYSMDKLELLG 60
OY 55 TMAAIIHGALPLMLVGNMTDSFA-NAGISRNKTFPVIIINESITNTTOHFINHLEEM 113
DB 61 TMAIAHSGGLPLMIVTGEHTDKFVNAG--NFSLPVNSLSMIMGR-----ILEEEM 113
OY 114 TTYAAYYSIGAGLVAAVYIOVSFWCLAGQOLIKRKOPFHAIIMROEIGWVHYGEL 173
DB 114 TTYAAYYSIGAGLVAAVYIOVSFWTLAAGROIKKIRONFHAIRLOREMGEFDIKGTTEL 173
OY 174 NTRLDDVSKINIEGDKIGMFHSHIAFEFTGIVGFIRGKLLVLIVLAIISPVLGSAAI 233
DB 174 NTRLDDVSKISIEGIGDVGMFQVAVAFGIVGFLIRGKLLVLIVLAIISPVLGSAAV 233
OY 234 WKIIISFDKRLAVAKAGAAVEVLAIKRTVIAFGQKKELBRYNKNLEAKGIGIKK 293

DB 234 WKIIISFDKRLAVAKAGAAVEVLAIKRTVIAFGQKKELBRYNKNLEAKGIGIKK 293
OY 294 AITANISIGAAFLIIVASYALAEVYISLVSESYISGIVLYTFVFSVLIAFSIGASPS 353
DB 294 AITANISMGIAFLIIVASYALAEVYISLVSESYISGIVLYTFVFSVLIAFSIGASPS 353
OY 354 IAFANARAAAEIKIIDNPSIDSYSGSKGKPDNKGLEFNKVFHSPSRKEVILK 413
DB 354 IAFANARAAAEIKIIDNPSIDSYSGSKGKPDNKGLEFNKVFHSPSRKEVILK 413
OY 414 GNLKVGSGQVYALVNGSGCKSTVQLMQRLLDPTDGMVCIQDQIRITINVRILRITG 473
DB 414 GNLKVGSGQVYALVNGSGCKSTVQLMQRLLDPTDGMVCIQDQIRITINVRILRITG 473
OY 474 VVSQEPVLEATTIAENIRKRGRENTMDLEKAVKANAVIDFMKLPKPFVLVGERGAQL 533
DB 474 VVSQEPVLEATTIAENIRKRGRENTMDLEKAVKANAVIDFMKLPKPFVLVGERGAQL 533
OY 534 SGQKORIAIARLVNPKIILLDEATSLDTESEAVVQVALDKARKGRTTVIAHRLST 593
DB 534 SGQKORIAIARLVNPKIILLDEATSLDTESEAVVQVALDKARKGRTTVIAHRLST 593
OY 594 VRNADYIAGFDGVIYKGNHDELMKEGITYFKLYTQTRNEIELENAATESSESAL 653
DB 594 VRNADYIAGFDGVIYKGNHDELMKEGITYFKLYTQTRNEIELENAATESSESAL 653
OY 654 EMSPKSGSLIKRSTRSHAPQGDRLKTKED-LNENAPVPSFMRILKLNSTEMPY 712
DB 654 GMP-NGWMSHIFRNSTKSLSSRAHHRLVDYDELDAVNPVPSFLKVLKLNSTEMPY 712
OY 713 FVVGIFCAIINGSLQPAFISITSRIGITFTRDEDPETKRQNSNMFSVLFVLGITSFTF 772
DB 713 FVVGITVAIVNGALQPAISITSEMIATFPGDD-AVKQOKNLFSLFVLGITSFTF 771
OY 773 FLOGFTGKAGELLTKRLRWYRSMLRDVSMFPDQKMTGALTTRLANDAAVQKGAIG 832
DB 773 FLOGFTGKAGELLTKRLRWYRSMLRDVSMFPDQKMTGALTTRLANDAAVQKGAIG 831
OY 833 SRLAVIQNTANIGTGIISLIYGMQTLTLLAIYPIIAIAGVEMKMLSGQALDKKEL 892
DB 833 TRLALIONTANIGTGIISLIYGMQTLTLLAIYPIIAIAGVEMKMLSGQALDKKEL 891
OY 893 EGAKRIATEALENFRVVSILTRQKFEWTAQSLQVPIYNSLRKHAIRFVSISIQAMV 952
DB 893 EAQKRIATEALENFRVVSILTRQKFEWTAQSLQVPIYNSLRKHAIRFVSISIQAMV 951
OY 953 FSYAGCRRFGAYIVANEFNMPDVLVFSATVYGAAVGVSSFAVDYAKAKVSAHVIM 1012
DB 953 FSYAGCRRFGAYIVANEFNMPDVLVFSATVYGAAVGVSSFAVDYAKAKVSAHVIM 1011
OY 1013 IIEKSPILDSYSPGLKPNLTLEGNVTENEVENYPTRPDIPVYLQIGLSLVKKGQTLALVG 1072
DB 1013 IIEKSPILDSYSPGLKPNLTLEGNVTENEVENYPTRPDIPVYLQIGLSLVKKGQTLALVG 1071
OY 1072 SSGGCKSTVQVOLLRRFPDPLAGSVLIDGKEIKHLNVOMLRANHGIYSQPIILFDCSIAEN 1132
DB 1072 SSGGCKSTVQVOLLRRFPDPLAGSVLIDGKEIKHLNVOMLRANHGIYSQPIILFDCSIAEN 1131
OY 1132 IAYGDSNKNVQODELVRAAKAANHPLFETLPQKYKTRVGDGTQSLSGGQKRLAIRRL 1191
DB 1132 IAYGDSNKNVQODELVRAAKAANHPLFETLPQKYKTRVGDGTQSLSGGQKRLAIRRL 1191
OY 1193 VROPHILLDEATSLDTESEKVVQVALDKARREGRTCIYIAHRLSTIONADLLIVYQNGK 1252
DB 1193 VROPHILLDEATSLDTESEKVVQVALDKARREGRTCIYIAHRLSTIONADLLIVYQNGK 1251
OY 1252 VKEHGTQOOLAQKCIYFSMVSVOAGAK 1280
DB 1252 VKEHGTQOOLAQKCIYFSMVSVOAGAK 1279
RESULT 10

Db	61	ALAHSGCLPLMIAMTVFGBMDKFDVNNAG--NFSLPVNFSLMLNBNGR----ILBEEMTRY	1133
QY	117	AYYSGAGVLAAYIAVYSFWCLAAGRQILKIRKOFPAIRROBIEIGFVDHVGELNTR	176
Db	114	AYYSGIGGVLLAAVIOYSFWTLAAGROIRKIRORFFHALLRQMGHFDIKGTTELNTR	1737
QY	177	LTDVSKINGIDCKIGMFHSLATFFGFIYGFTRKMLTVLIATSPVIGLSAAIMAK	236
Db	174	LTDISKISISGIDCKGMFQOALTEFFAGFIVGFIRGKMLTVLIMATVAILGISTAVMAK	233
QY	237	ILSFSDKELLAAKAGAAVEELAIRVIVFGOKKELEYNKNLNEAAGTICKAIT	296
Db	234	ILSFSDKELLAAKAGAAVEELAIRVIVFGQNKLEERYOKHLENNAKTIKRAIS	293
QY	297	ANISGAPELLIYASTALAFWYGTSLVSSEXYSIGOVLTFEFSYLGAISGAQSPSIEA	356
Db	294	ANISMIAELLIYASTALAFWYGTSLVLISETTYIGNATVFFSILIGAFVGOQAPOIDA	353
QY	357	FANARAAAEIFRIIDNKSIDSYSKSGHKDNIKGNLEFNHFSYPSREKVIILGLN	416
Db	354	FANARAAAVIETDIIDNNPKIDSFSEGRKHPDSIKGNLEFSDVHFSPSRANKILKGLN	413
QY	417	LKVSQGVVALVNSGCCSTVTQLMQRILYPTDGMVCIDQDRIITNVRLHREITGVY	476
Db	414	LKVSQGVVALVNSGCCSTVTQLOLRDYPTEGTSISDQDRIIRNNVNCLEHFGIVS	473
QY	477	QEBVLEFATIANINISYGRNATMDIEIKAYVEANNAFYIKLPMKEDTLVGERAOISGC	536
Db	474	QEBVLESTIANINIRKGRNATMDIEIKAYVEANNAFYIMKLPKEDTLVGDGAOLSG	533
QY	537	OKORIATAALVRNKRILLDEATSAIDTESEAAVOVALDKARGRITVIYARLSVTR	596
Db	534	OKORIATAALVRNPKILLDEATSAIDTESEARVOAALDKARGRTVIYARLSVTR	593
QY	597	ADVIAGFDQGVIVEGNDELMEKEGYIFKLYVMQTRGNETIELENATGESKSSDALEMS	656
Db	594	ADVIAGFEGDGVIVEGSHSELIRKEGYIFRLVNNQTSOILSEEFLEVELSDKAAGVA	653
QY	657	PDKSSSLIKRSTRSRSHAHPOGDRIKGT-EDLNANPVPSFWRLIKNSTEMEYFV	715
Db	654	P-NGMKARIFRNSIKTSLKSSRAHONRLDVTENLDANVPVSFLKYLKRNKTEWYFV	712
QY	716	GIFCALINGLOPAFSIIFRSIIIGITFREDDEPSTRONSNMFVLEVLGIIISIFFFLO	775
Db	713	GLCLAIANGALOPAFSIILSEMIAIFEGBD-TVQOKCNMFSLVFLGIVHSFFFFILO	771
QY	776	GTFEKAEBIILTRLRVYVPSMRROOVSPFDOPKNTGMLTTRLANDAAOYKGAIGSL	835
Db	772	GTFEKAEBIILTRLSRAEFAAMLRQMSFMDHKNSTGALSTRLATDAOYOGATGRIL	831
QY	836	AVITONIANIGTGIISLIVGOWJTLLILVPIALIAAGVEKMLSGOALDKRKELEBA	895
Db	832	ALIANONTANIGTGIISPIYGMQWJTLLLSVFEFVAVAGIVEKMLAGNAKDKKEMEA	891
QY	896	GKIAEALENRIVYSLTRECKFEYMAQSOLOVPRNSLKAHIFGVSVSTIQAMNYSY	955
Db	892	GKIAEALENRIVYSLTOERKESMEVKEILHOPRNSYAKAHIIYGFISIQAMNYSY	951
QY	956	AGCFRFGAYIVLANEFMNFQVLLVESAIYVGAAGVOSSPADYAKAVSAAHIMITE	1015
Db	952	AGCFRFGSLVIVNHFMRKQVILVFSAIYVGAVALDHASSPADYAKAKLASAAYLSPE	1011
QY	1016	KSPILIDSTPHGLKPNLTBSNVTNFVEVNYPTRPDIPIVLOGLSLEVKKGQTLALVSSG	1075
Db	1012	RQPLIDSTSRGMPDKFESGVTFNEVNYPTRAWVPIVLOGLSLEVKKGQTLALVSSG	1071
QY	1076	CGKSTVQVLLERFDPPLAGVLLIDGKEIKHIANOMRAHLGIYVSEPIITFQSTANENAY	1135
Db	1072	CGKSTVQVLLERFDPMAAGVLLIDGQPAKKLANQMLRAOLGIYVSEPIITFQSTAKNATY	1131
QY	1136	GDNSRVVSHETMOAKKANIHFIETLDEKYNTRVGDKGTOLSGQOKORIATAALVRQ	1195
Db	1132	GDNSRVVSHETMOAKKANIHFIETLDEKYNTRVGDKGTOLSGQOKORIATAALVRQ	1191

Oy	1196	PHILIDEATSAIDTSESEVVOEALDKRAREGTCIVIAHRLSTIONANDLIVYQNKKE	1255
Dd	1192	PRVLIDAEATSADTSESEVVOEALDKRAREGTCIVIAHRLSTIONANDLIVYQNKKE	1251
Oy	1256	HGHMOQLAOKGYIFSMVSVOAGAK	1280
Dd	1252	HGHMOQLAOKGYIFSMVINQAGTO	1276
 RESULT 11			
AB11	HUMAN	STANDARD:	PRT; 1321 AA.
AC	095342;	09UNB2;	
DT	16-OCT-2001	(Rel. 40,	Created)
DT	16-OCT-2001	(Rel. 40,	Last sequence update)
DT	15-JUN-2002	(Rel. 41,	Last annotation update)
DE	Bile salt export pump	(ATP-binding cassette, sub-family B, member 11).	
GN	ABCBI1 OR BSPP.		
OS	Homo sapiens (Human).		
OC	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;		
NCBI_TaxID=9606;	Mammalia; Eutheria; Primates; Catarrhini; Hominoidea; Homo.		
RN	SEQUENCE FROM N.A., AND VARIANTS PFC2 GLY-297; GLU-461; GLY-482;		
RX	ARG-982; CYS-1153 AND GLU-1268.		
RX	MEDLINE=99021377; PubMed=9806540;		
RA	Stratniels S.S., Bull L.N., Katsely A.S., Kocoshis S.A., Dahl N.,		
RA	Arneil H., Sokal E., Dahan K., Childs S., Ling V., Tanner M.S.,		
RA	Kargwalia A.F., Nemeth A., Pavlovskaja J., Baker A., Mell-Vergant G.,		
RA	Frelmer N.B., Gardiner R.M., Thompson R.J.;		
RT	"A gene encoding a liver-specific ABC transporter is mutated in		
RT	progressive familial intrahepatic cholestasis.";		
RT	Nat. Genet. 20:233-238(1998).		
RN	[2]		
RN	SEQUENCE FROM N.A.		
RA	Mol O., Hoofveld G.J.E.J., Jansen P.L.M., Muller M.;		
RT	"Cellular localization and functional characterization of the human		
RT	bile salt export pump (BSPP).";		
RT	Submitted (MAR-1999) to the EMBL/Genbank/DDBJ databases.		
RN	[3]		
RX	VARIANT ALA-444.		
RX	MEDLINE=21686803; PubMed=11829140;		
RA	Saito S., Iida A., Sekine A., Miura Y., Ogawa C., Kawachi S.,		
RA	Higuchi S., Nakamura Y.;		
RT	"Three hundred twenty-six genetic variations in genes encoding nine		
RT	members of ATP-binding cassette, subfamily B (ABCB/MDR/TAP), in the		
RT	Japanese population.";		
RT	J. Hum. Genet. 47:38-50(2002).		
RN	[4]		
RN	VARIANTS INTRAHEPATIC CHOLESTASIS LEU-284 AND ASP-1004.		
RX	MEDLINE=21674594; PubMed=11815775;		
RA	Chen H.-L., Chang P.-S., Hsu H.-C., Ni Y.-H., Hsu H.-Y., Lee J.-H.,		
RA	Jeng Y.-M., Shau W.-Y., Hsu M.-H.;		
RT	"PFC1 and BSP defects in Taiwanese patients with chronic intrahepatic		
RT	cholestasis with low gamma-glutamyltranspeptidase levels.";		
RT	J. Pediatr. 140:119-124(2002).		
CC	-1- FUNCTION: INVOLVED IN THE ATP-DEPENDENT SECRETION OF BILE SALTS		
CC	INTO THE CANALICULUS OF HEPATOCTES.		
CC	-1- SUBCELLULAR LOCATION: Integral membrane protein.		
CC	-1- TISSUE SPECIFICITY: EXPRESSED PREDOMINANTLY IN		
CC	THE LIVER, WHERE IT WAS FURTHER LOCALIZED TO THE CANALICULAR		
CC	MICROVILLI AND TO SUBCANALICULAR VESICLES OF THE HEPATOCTES BY		
CC	SITO (BY SIMILARITY).		
CC	-1- DOMAIN: MULTIFUNCTIONAL POLYPEPTIDE WITH TWO HOMOLOGOUS HALVES,		
CC	EACH CONTAINING AN HYDROPHOBIC MEMBRANE-ANCHORING DOMAIN AND AN		
CC	ATP BINDING CASSETTE (ABC) DOMAIN.		
CC	-1- DISEASE: DEFECTS IN ABCBI1 ARE THE CAUSE OF PROGRESSIVE FAMILIAL		
CC	INTRAHEPATIC CHOLESTASIS 2 (PFC2), AN INHERITED LIVER DISEASE OF		
CC	CHILDHOOD. PFC2 IS CHARACTERIZED BY CHOLESTASIS AND NORMAL SERUM		
CC	GAMMA-GUTAMYLTRANSFERASE ACTIVITY. DEFECTS IN ABCBI1 ARE ALSO		
CC	FOUND IN CASES OF CHRONIC INTRAHEPATIC CHOLESTASIS WITHOUT OBVIOUS		
CC	FAMILIAL HISTORY OF CHRONIC LIVER DISEASE.		

CC	-I- SIMILARITY: BELONGS TO THE ABC TRANSPORTER FAMILY. MDR SUBFAMILY.
CC	-----
CC	This SWISS-PROT entry is copyright. It is produced through a collaboration
CC	between the Swiss Institute of Bioinformatics and the EMBL Outstation -
CC	the European Bioinformatics Institute. There are no restrictions on its
CC	use by non-profit institutions as long as its content is in no way
CC	modified and this statement is not removed. Usage by and for commercial
CC	entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC	or send an email to license@sib.ch).
CC	-----
DR	EMBL; AF091582; AAC77455.1; .
DR	EMBL; AE136523; AAD28285.1; .
DR	Genew: HGNC:42; ABCB11.
DR	MIM; 603201; .
DR	MIM; 601847; .
DR	InterPro: IPR003593; AAA_ATPase.
DR	InterPro: IPR003439; ABC_transporter.
DR	InterPro: IPR001140; Abctransplrtm.
DR	Pfam: PF000005; ABC_tran; 2.
DR	Pfam: PF00664; ABC_membrane; 2.
DR	ProDom: PD000006; ABC_transportr; 2.
DR	SMART; SM00382; AAA; 2.
DR	PROSITE; PS00211; ABC_TRANSPORTER; 1.
KW	ATP-binding; Transmembrane; transport; Disease mutation; Polymorphism.
FT	DOMAIN 1 62
FT	TRANSMEM 63 83
FT	DOMAIN 84 147
FT	TRANSMEM 148 168
FT	DOMAIN 169 215
FT	TRANSMEM 216 236
FT	DOMAIN 237 240
FT	TRANSMEM 241 261
FT	DOMAIN 262 319
FT	TRANSMEM 320 340
FT	DOMAIN 341 353
FT	TRANSMEM 354 374
FT	DOMAIN 375 755
FT	TRANSMEM 756 776
FT	DOMAIN 777 794
FT	TRANSMEM 795 815
FT	DOMAIN 816 869
FT	TRANSMEM 870 890
FT	TRANSMEM 891 911
FT	DOMAIN 912 979
FT	TRANSMEM 980 1000
FT	DOMAIN 1001 1011
FT	TRANSMEM 1012 1032
FT	DOMAIN 1033 1321
FT	NP_BIND 455 462
FT	NP_BIND 1113 1120
FT	CARBOHYD 109 109
FT	CARBOHYD 116 116
FT	CARBOHYD 122 122
FT	CARBOHYD 125 125
FT	VARIANT 284
FT	-----
FT	VARIANT 297 297
FT	-----
FT	VARIANT 444 444
FT	-----
FT	VARIANT 461 461
FT	-----
FT	VARIANT 482 482
FT	-----
FT	VARIANT 982 982
FT	-----
FT	VARIANT 1004 1004
FT	-----
FT	VARIANT 1153 1153
FT	-----
FT	VARIANT 1268 1268
FT	-----
FT	R -> Q (IN PFC2).
FT	/FTId-VAR_013338.
FT	R -> C (IN PFC2).
FT	/FTId-VAR_013337.
FT	G -> R (IN PFC2).
FT	/FTId-VAR_013335.
FT	D -> G (IN PFC2).
FT	/FTId-VAR_013334.
FT	K -> E (IN PFC2).
FT	/FTId-VAR_013333.
FT	V -> A.
FT	/FTId-VAR_010271.
FT	E -> G (IN PFC2).
FT	/FTId-VAR_010271.
FT	CHOLESTASIS).
FT	/FTId-VAR_013336.
FT	G -> D (IN CHRONIC INTRAHEPATIC
FT	CHOLESTASIS).
FT	/FTId-VAR_013337.
FT	R -> C (IN PFC2).
FT	/FTId-VAR_013338.
FT	R -> Q (IN PFC2).
FT	/FTId-VAR_013339.
FT	R -> Q (IN PFC2).
FT	/FTId-VAR_013338.
FT	R -> Q (IN PFC2).
FT	/FTId-VAR_013337.
FT	R -> Q (IN PFC2).
FT	/FTId-VAR_013336.
FT	R -> Q (IN PFC2).
FT	/FTId-VAR_013335.
FT	R -> Q (IN PFC2).
FT	/FTId-VAR_013334.
FT	R -> Q (IN PFC2).
FT	/FTId-VAR_013333.
FT	R -> Q (IN PFC2).
FT	/FTId-VAR_013332.
FT	R -> Q (IN PFC2).
FT	/FTId-VAR_013331.
FT	R -> Q (IN PFC2).
FT	/FTId-VAR_013330.
FT	R -> Q (IN PFC2).
FT	/FTId-VAR_013329.
FT	R -> Q (IN PFC2).
FT	/FTId-VAR_013328.
FT	R -> Q (IN PFC2).
FT	/FTId-VAR_013327.
FT	R -> Q (IN PFC2).
FT	/FTId-VAR_013326.
FT	R -> Q (IN PFC2).
FT	/FTId-VAR_013325.
FT	R -> Q (IN PFC2).
FT	/FTId-VAR_013324.
FT	R -> Q (IN PFC2).
FT	/FTId-VAR_013323.
FT	R -> Q (IN PFC2).
FT	/FTId-VAR_013322.
FT	R -> Q (IN PFC2).
FT	/FTId-VAR_013321.
FT	R -> Q (IN PFC2).
FT	/FTId-VAR_013320.
FT	R -> Q (IN PFC2).
FT	/FTId-VAR_013319.
FT	R -> Q (IN PFC2).
FT	/FTId-VAR_013318.
FT	R -> Q (IN PFC2).
FT	/FTId-VAR_013317.
FT	R -> Q (IN PFC2).
FT	/FTId-VAR_013316.
FT	R -> Q (IN PFC2).
FT	/FTId-VAR_013315.
FT	R -> Q (IN PFC2).
FT	/FTId-VAR_013314.
FT	R -> Q (IN PFC2).
FT	/FTId-VAR_013313.
FT	R -> Q (IN PFC2).
FT	/FTId-VAR_013312.
FT	R -> Q (IN PFC2).
FT	/FTId-VAR_013311.
FT	R -> Q (IN PFC2).
FT	/FTId-VAR_013310.
FT	R -> Q (IN P

```

FT  CONFLICT      339      339      /FTID=VAR_013339.
50  SEQUENCE      1321 AA; 146392 MW; D44CACCA8DEA371 CRC64;

Query Match
Best Local Similarity 50.3%; Pred. No. 7,8e-179;
Matches 657; Conservative 242; Mismatches 367; Indels 40; Gaps 10;

QY  8 KGAENKFWMGKRSKKNKK-----EKK---PTVSTFAMFYSNMIDRLVMTVMTMAI 59
    |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |
DB  11 KKEGEENDEPESDKSYNNKSKSLDQEKGDGVKRGVGFQDLRFSSSTDIWIMLFSCAF 70
    |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |
QY  60 IHGAALPLMALVYGNMTDFPANGISR-----NKTFF---PVINESITNTQHF 105
    |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |
DB  71 LHGIADPGVLLFGTWTDFEIDYDVELQIIPGKACVNMNTIVMNSLSNMNTNGTFCG 130
    |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |
QY  106 INHLEEMTYYAAYYSAGVLAAYIQVSWCLAAQROLIKRKQFPAHIMROETGMF 165
    |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |
DB  131 LNTIESEMKEFASYIAGIYAVLITGYIQCWVIAAARQIKMKKFFRRIMHEIGWF 190
    |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |
QY  166 DVHVDGELNTRLDVSKINEGIDGKIGMFESHIAFTFGTGVGTRGMKTLVILASP 225
    |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |
DB  191 DCSVVELNTRRESDDINKINDAIADQMALFIQRMNSTICGFLGEPFGMKTLVILISVP 250
    |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |
QY  226 VIGLSAIAKAKITLSTFDEKELATYAKAGVAEVLAIKIVIAFGGKKELERYNKNLEE 285
    |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |
DB  251 LIGIGAITIGLSVSKFTYELKAAKAGVADDEVISSMRTVAAPGKREVEREKNLVF 310
    |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |
QY  286 AAGIGIKKAITANISIGAFLIYAYALAFWYGSVLV--SSEKSTQVLTVEFSVLGA 344
    |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |
DB  311 AQRMGIRKGIYVGFGTGKWLICLVANATWYSTLYLDEGETTPTVLQVILSYVGA 370
    |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |
QY  345 PSIGASISIAFANAGAAVEIKFIIDNKPISIDYSKSGHKPDINKNLKFNKVPSP 404
    |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |
DB  371 LNLGASPCLEAFATGRAAATSIETIDRKPIIDCMSEDEGKIDRIKGEIEFNVPFHY 430
    |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |
QY  405 SRKEKFIKGLNKLKVGOSQOTATLVNNGSGCKSTVOQLMQLYPTDGMVTCODIRTN 464
    |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |
DB  431 SRPEKILINDNMVTKREEMALVGPAGAGSTALQILOFRPYDCRMVVDHDIRSN 490
    |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |
QY  465 VRIHREITGVVSOEPLVETATTIAENIRYGRBNVTMDEIEKAVEANAADPEIMLPKPT 524
    |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |
DB  491 IOWLMDQIGIYEOEPVLESTIAENIRYGRBDATMEDIYQAEEANAYNFIMDLPOQFT 550
    |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |
QY  525 LVGEGAGQSGGOKORIAALALVNRKILLDBAISALDTESEAVVOVALDKARKGRT 584
    |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |
DB  551 LVGEGGOMSGGOKORVAIALALRNKILLDMATSDLDNESEAAVQEVLSKHGHTI 610
    |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |
QY  585 IVIAHRLSTVNADVIAGFDGVIVKGNHDELMKEGIFYFKLYMTQGTG---NEIELE 640
    |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |
DB  611 ISVAHRLSTVRAADTIIGFEHGTAVEGHTEHLEKRGVITLVTLQSGONALNEEDIK 670
    |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |
QY  641 NATGESSESDALEMPKDGSSILKRSTR--SIHAP--OGODRKGTGED----- 689
    |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |
DB  671 DATDDMALARTFSGSYDILRASIRORSKQSLSYLVHPEPLAVVDHKTSTYEDKMDI 730
    |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |
QY  690 -LANENPVVPSFWMRLKLNSTEMPYFVWGIFCAIINGLOPASIIFSRITGIFTDDEDE 748
    |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |
DB  731 PVGEVEVPAVRRIKLFSAPEWPMYLVSGAVAVNCTVPLPAFLFSQILGTFSS-IPKPE 789
    |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |
QY  749 TKRONSMESVFLVLIIGISITFFLQGTFFGKAGEILIKRLRYMFRSMLEQDVSFDP 808
    |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |
DB  790 -EORSQINGVCLLFVAMGCVSLFQFLOGYAFKSGELLLKRRKKGCFRAMLGODIAMDD 849
    |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |
QY  809 PKNTGTALTTRLANDAAOVKGAIGSLAVITONIANLGTGIISILYQMQLTLLLATVP 868
    |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |
DB  850 LRSPGALTRLATDASQVGAAGSQIGMIVSFTNVAMIAFSPFWMKSLIVLCFP 909
    |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |
QY  869 IIAIAGVEMKMLSGQALDKKELEGAKIATEAENFTVYSLRREKGFYMAQSLQV 928
    |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |
DB  910 FLALSGATQTRMLTGAFASDKQALEWVGQITTEALSNIRYVAGIKERKRFTEALETELEK 969
    |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |
QY  929 PYNSLRKRIHIGVSRSTIQAMMYFSYAGCERFAGYLVANEFMNOVDVLLVSAIVEGAM 988

```

```

DB  970 PFTALQKANIYGFCEFAQOCINFANSASYRYGGILLISNEGILHFSYFVIAVLSAT 1029
    |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |
QY  989 AVGVSSFADPYAKAKVYSAAHVIMIEKSPLDYSFPHLKNTLEGVNTFVENVFPT 1048
    |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |
DB  1030 ALGRATSYTPSYAKAKISARFFQQLDRQPPISVNTVAGEKMDNQGIIDFVDCKFYPS 1089
    |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |
QY  1049 RPDIPVLOQISLEVRKQGTALVGVSSGCKSTVOLLERFYPLAGSLVIGKEIKHNV 1108
    |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |
DB  1090 RPDQVNLNGISVSIISPQTLARFVSSGCKSTIDLERFDPDQKMAIDGHSKAYNV 1149
    |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |
QY  1109 QMIRAHILYVSEPILEFDSIAENTAYGDNRSVYSHHEIMQAAKANIHETLPEKYN 1168
    |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |
DB  1150 QPLRSNIGVISOEPLVFACSINDNIKYGDKTEIPMERVIAAARQALHDFVMSLPEKE 1209
    |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |
QY  1169 TRVGDKTOLSGGOKORIAIARALVROPHILLDBATSALDTESEKVVQELDKARERT 1228
    |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |
DB  1210 TNGVSGGQSLSGEKORIAIAIYRDKILLDBATSALDTESEKTVQVALDKARERT 1269
    |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |
QY  1229 CIVIAHRLSTIONADLIIVFPONGKVEKGTQOQLAKGIFYSMWS 1274
    |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |
DB  1270 CIVIAHRLSTIONADIIIVMAQGVIEKGTHEIMAQAGAYKLVLT 1315

```

RESULT 12

```

AB1_RABIT
ID AB1_RABIT STANDARD; PRT; 1321 AA.
AC 09NOV3;
AD 16-OCT-2001 (Rel. 40, Created)
AD 16-OCT-2001 (Rel. 40, Last sequence update)
AD 16-OCT-2001 (Rel. 40, Last annotation update)
DE Bile salt export pump (ATP-binding cassette, sub-family B, member 11)
DE (Sister of P-glycoprotein).
GN ABCB11 OR BSEP OR SPGP.
OS Oryctolagus cuniculus (Rabbit).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OM Mammalia; Eutheria; Lagomorpha; Leporidae; Oryctolagus.
OX NCBI_TaxID=9986;
RN [1]
RP
RC STRAIN=New Zealand white; TISSUE=Liver;
RA Balasubramanian N.V., Suchy F.J., Ananthanarayanan M.;
RT "Molecular cloning and characterization of rabbit liver bile salt
  export pump (Bsep/spgp).";
RL Submitted (MAR-2000) to the EMBL/GenBank/DBJ databases.
CC
CC -1- FUNCTION: INVOLVED IN THE ATP-DEPENDENT SECRETION OF BILE SALTS
  INTO THE CANALICULUS OF HEPATOCYTES.
CC
CC -1- SUBCELLULAR LOCATION: Integral membrane protein.
CC
CC -1- TISSUE SPECIFICITY: EXPRESSED PREDOMINANTLY, IF NOT EXCLUSIVELY IN
  THE LIVER, WHERE IT WAS FURTHER LOCALIZED TO THE CANALICULAR
  MICROVILLI AND TO SUBCANALICULAR VESICLES OF THE HEPATOCYTES BY IN
  SITU.
CC
CC -1- DOMAIN: MULTIFUNCTIONAL POLYPEPTIDE WITH TWO HOMOLOGOUS HALVES,
  EACH CONTAINING AN HYDROPHOBIC MEMBRANE-ANCHORING DOMAIN AND AN
  ATP BINDING CASSETTE (ABC) DOMAIN.
CC
CC -1- SIMILARITY: BELONGS TO THE ABC TRANSPORTER FAMILY. MDR SUBFAMILY.
CC
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
  between the Swiss Institute of Bioinformatics and the EMBL Outstation -
  the European Bioinformatics Institute. There are no restrictions on its
  use by non-profit institutions as long as its content is in no way
  modified and this statement is not removed. Usage by and for commercial
  entities requires a license agreement (see http://www.1sb-sib.ch/announce/
  or send an email to license@sib-sib.ch).
CC
EMBL: AF249879; AAF6552.1; -
DR InterPro: IPR003593; AAA_Atpase.
DR InterPro: IPR003439; ABC_transporter.
DR InterPro: IPR001140; AbctransprtTM.
DR Pfam: PF00005; ABC_tran.
DR Pfam: PF00664; ABC_membrane.
DR ProDom: PD000006; ABC_transportr; 2.
DR SMART: SM00382; AAA; 2.

```

Query Match	49.88;	Score 3224.5;	DB 1;	Length 1321;
Best Local Similarity	49.48;	Pred. NO. 1.4e-174;		
Matches 638;	Conservative 249;	Mismatches 363;	Indels 41;	Gaps 8;

Accession	Gene	Species	Protein	Length	MD5	SHA1	SHA256	SHA512	SHA1024	SHA2048	SHA4096	SHA8192	SHA16384	SHA32768	SHA65536	SHA131072	SHA262144	SHA524288	SHA1048576	SHA2097152	SHA4194304	SHA8388608	SHA16777216	SHA33554432	SHA67108864	SHA134217728	SHA268435456	SHA536870912	SHA1073741824	SHA2147483648	SHA4294967296	SHA8589934592	SHA17179869184	SHA34359738368	SHA68719476736	SHA137438953472	SHA274877906944	SHA549755813888	SHA1099511627776	SHA2199023255552	SHA4398046511104	SHA8796093022208	SHA1759218644416	SHA3518437288832	SHA7036874577664	SHA14073749155328	SHA28147498310656	SHA56294996621312	SHA112589993242624	SHA225179986485248	SHA45035997297056	SHA90071994594112	SHA180143989188224	SHA360287978376448	SHA720575956752896	SHA1441151913511776	SHA2882303827023552	SHA5764607654047104	SHA11529215308094208	SHA23058430616188416	SHA46116861232376832	SHA92233722464753664	SHA184467444929507200	SHA36893488985901440	SHA73786977971802880	SHA147573955943605760	SHA295147911887211520	SHA590295823774423040	SHA1180591647548846080	SHA2361183295097692160	SHA4722366590195384320	SHA9444733180390768640	SHA1888946636078133280	SHA3777893272156266560	SHA7555786544312533120	SHA15111573088625066240	SHA30223146173250132480	SHA60446292346500264960	SHA120892584693000529920	SHA241785169386001059840	SHA483570338772002119680	SHA967140677544004239360	SHA1934281355088008478720	SHA3868562710176016957440	SHA7737125420352033914880	SHA15474250840704068297600	SHA30948501681408136595200	SHA61897003362816273190400	SHA12379400672573254638080	SHA24758801345146509276160	SHA49517602690293018553280	SHA99035205380586037106560	SHA19807041076117207421120	SHA39614082152234414842240	SHA79228164304468829684480	SHA158456328608937659368960	SHA316912657217875318737920	SHA633825314435750637475840	SHA1267650628871501274951680	SHA2535301257743002549903360	SHA5070602515486005099806720	SHA10141205030972010199613440	SHA20282410061954020399226880	SHA40564820123908040798453760	SHA81129640247816081596887520	SHA16225928049562116317377280	SHA32451856099124232634755360	SHA64903712198248465269510720	SHA129807424396488905339021440	SHA259614848792977810678042880	SHA51922969758595562137685760	SHA10384593951719112227537153280	SHA2076918790343822445507426560	SHA415383758068764489101481120	SHA830767516137528978202882240	SHA166153503227505796405764480	SHA3323070064550115928115288960	SHA6646140129100231856230577920	SHA13292280258000437124611557440	SHA26584560516000874252430110880	SHA53169121032001748504860221760	SHA10633824206400349009721243520	SHA21267648412800698019442487040	SHA42535296825601396038884974080	SHA850705936512027920777698880	SHA1701411873024058401555377760	SHA340282374604811680311075553280	SHA680564749209623360622151106560	SHA136112949841934672444222213120	SHA27222589968386934884884442240	SHA54445179936773869769768884480	SHA10889035975554773953953776960	SHA2177807195111154790790715553280	SHA435561439022230958181831106560	SHA87112287804446191736363622113120	SHA1742245566088833834672726222240	SHA3484491132177766773344444444480	SHA6968982263555533446688888888960	SHA13937964527111111111111111111040	SHA27875929054222222222222222222080	SHA55751858108444444444444444444160	SHA11150371621688888888888888888320	SHA22300742432177777777777777777640	SHA44601484864355555555555555555280	SHA8920296972871111111111111111111040	SHA1784059395574222222222222222222080	SHA3568118791148444444444444444444160	SHA71362375822888888888888888888320	SHA14272475165777777777777777777640	SHA28544951321555555555555555555280	SHA5708990264311111111111111111111040	SHA1141798052862222222222222222222080	SHA2283596112424444444444444444444160	SHA4567192224848888888888888888888320
-----------	------	---------	---------	--------	-----	------	--------	--------	---------	---------	---------	---------	----------	----------	----------	-----------	-----------	-----------	------------	------------	------------	------------	-------------	-------------	-------------	--------------	--------------	--------------	---------------	---------------	---------------	---------------	----------------	----------------	----------------	-----------------	-----------------	-----------------	------------------	------------------	------------------	------------------	------------------	------------------	------------------	-------------------	-------------------	-------------------	--------------------	--------------------	-------------------	-------------------	--------------------	--------------------	--------------------	---------------------	---------------------	---------------------	----------------------	----------------------	----------------------	----------------------	-----------------------	----------------------	----------------------	-----------------------	-----------------------	-----------------------	------------------------	------------------------	------------------------	------------------------	------------------------	------------------------	------------------------	-------------------------	-------------------------	-------------------------	--------------------------	--------------------------	--------------------------	--------------------------	---------------------------	---------------------------	---------------------------	----------------------------	----------------------------	----------------------------	----------------------------	----------------------------	----------------------------	----------------------------	----------------------------	----------------------------	----------------------------	-----------------------------	-----------------------------	-----------------------------	------------------------------	------------------------------	------------------------------	-------------------------------	-------------------------------	-------------------------------	-------------------------------	-------------------------------	-------------------------------	-------------------------------	--------------------------------	--------------------------------	-------------------------------	----------------------------------	---------------------------------	--------------------------------	--------------------------------	--------------------------------	---------------------------------	---------------------------------	----------------------------------	----------------------------------	----------------------------------	----------------------------------	----------------------------------	----------------------------------	--------------------------------	---------------------------------	-----------------------------------	-----------------------------------	-----------------------------------	----------------------------------	----------------------------------	----------------------------------	------------------------------------	-----------------------------------	-------------------------------------	------------------------------------	------------------------------------	------------------------------------	-------------------------------------	-------------------------------------	-------------------------------------	-------------------------------------	-------------------------------------	-------------------------------------	---------------------------------------	---------------------------------------	---------------------------------------	-------------------------------------	-------------------------------------	-------------------------------------	---------------------------------------	---------------------------------------	---------------------------------------	---------------------------------------

RT "Molecular cloning and characterization of the murine bile salt export
 RT pump.";
 RL Gene 241:117-123(2000).
 RN [2]
 RP SEQUENCE OF 463-635 FROM N.A.
 RC STRAIN-C57BL/6; TISSUE-Liver:
 RA Salkar R., Suchy F.J., Ananthanarayanan M.;
 RT "Molecular cloning of mouse liver bile salt export pump (bsep).";
 RL Submitted (SEP-1999) to the EMBL/Genbank/DBP databases.
 CC -1- FUNCTION: INVOLVED IN THE ATP-DEPENDENT SECRETION OF BILE SALTS
 CC INTO THE CANALICULUS OF HEPATOCYTES.
 CC -1- SUBCELLULAR LOCATION: Integral membrane protein.
 CC -1- TISSUE SPECIFICITY: EXPRESSED PREDOMINANTLY, IF NOT EXCLUSIVELY IN
 CC THE LIVER, WHERE IT WAS FURTHER LOCALIZED TO THE CANALICULAR
 CC MICROVILLI AND TO SUBCANALICULAR VESICLES OF THE HEPATOCYTES BY IN
 CC SITU.
 CC -1- DOMAIN: MULTIFUNCTIONAL POLYPEPTIDE WITH TWO HOMOLOGOUS HALVES,
 CC EACH CONTAINING AN HYDROPHOBIC MEMBRANE-ANCHORING DOMAIN AND AN
 CC -1- SIMILARITY: BELONGS TO THE ABC TRANSPORTER FAMILY. MDR SUBFAMILY.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL Outstation-
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC -----
 DR EMBL: AF133903; AAF14372.1; -
 DR EMBL: AF186585; AAD56419.1; -
 DR MGI: MGI:131619; Abcb11.
 DR InterPro: IPR003593; AAA_ATPase.
 DR InterPro: IPR003439; ABC_transporter.
 DR Pfam: PF00005; ABC_tran; 2.
 DR Pfam: PF00664; ABC_membrane; 2.
 DR ProDom: PD000006; ABC_transport; 2.
 DR SMART: SM00382; AAA; 2.
 DR PROSITE: PS00211; ABC_TRANSPORTER; 1.
 DR KW ATP-binding; Transmembrane; Transport.
 FT DOMAIN 1 62
 FT TRANSMEM 63 83
 FT DOMAIN 84 147
 FT TRANSMEM 148 168
 FT DOMAIN 169 215
 FT TRANSMEM 216 236
 FT DOMAIN 237 240
 FT TRANSMEM 241 261
 FT DOMAIN 262 319
 FT TRANSMEM 320 340
 FT DOMAIN 341 353
 FT TRANSMEM 354 374
 FT DOMAIN 375 755
 FT TRANSMEM 756 776
 FT DOMAIN 777 794
 FT TRANSMEM 795 815
 FT DOMAIN 816 869
 FT TRANSMEM 870 890
 FT DOMAIN 891 911
 FT TRANSMEM 912 979
 FT DOMAIN 980 1000
 FT TRANSMEM 1001 1011
 FT DOMAIN 1012 1032
 FT TRANSMEM 1033 1321
 FT NP_BIND 455 462
 FT CARBOHYD 1113 1120
 FT CARBOHYD 109 109
 FT CARBOHYD 116 116
 FT CARBOHYD 122 122
 FT CARBOHYD 125 125
 FT CONFLICT 481 481
 FT CONFLICT 633 633

SQ SEQUENCE 1321 AA: 146675 MW: 1585E6F175D32967 CRC64;
 Query Match 49.5%; Score 3207.5; DB 1; Length 1321;
 Best Local Similarity 49.7%; Pred. No. 1,36-173;
 Matches 642; Conservative 237; Mismatches 377; Indels 37; Gaps 9;
 QY 15 FMKMGKSK-KNEKKKKPYSTFMEFYSNMDRLTYLVTGMAATIGALPLMLVFG 73
 Db 25 FHNNDKSKSLQDKKKDEGARVGFELFRSSSKDNMTLMFGSVICALDHGAQGMILIVG 84
 QY 74 NMDFPANAIGSR-----NKFPPV---INESINNNQPHFINLHEEMTYANY 119
 Db 85 ILTDFEVEDIERQRLSLIPKVCNNNTIWNISFNQNMNTNGISGIVDINSEVIGSGI 144
 QY 120 YSGIAGVLAAYVAYOVSWCLAAGRQLTKRKQFPAIMRQEGMDVDHDELMTRLTD 179
 Db 145 YAGVAVANLITGFOIRLWVTGARGQIRKMRKFFRIMMEIGWPDCTSVGEIINSRSD 204
 QY 180 DVSKINIEGIDGKIMFPHFSIATFEPTGIVGFTGKWLTLVITLAIISPYGLSAAIMAKILS 239
 Db 205 DINKIDEAIDOMALFLQRLSTALSGILGFTYKMKLITLIVASPLIGIGAAVIGLSVA 264
 QY 240 SFTDKELAYAKAGAVAEVLAIRTYIAFGGQKELERYNNLEBANGIGIKKAITANI 299
 Db 265 KFTLELKAIAKAGSIADDEVLSIRTYIAFGGENKEVEREYERKNLMFAQRWGIKGMVWGF 324
 QY 300 SIGAFLITVIAAYALAFWYGTSLV--SSEYSIGOVLTVEFVSILGAFSIGASISIAEFA 358
 Db 325 FTGYMCLIFCYVALAFWYSGSLVLDGEGYTGTLQIFLCYLIAMNIGNASSCLEIFS 384
 QY 359 NARGAAVEIKIIDNKPISDYSKSGHKPDNIKMLEKNVHFSPSRKEVILKGLNLK 418
 Db 385 TGCASASSIFQTLRQPVMDCMGSDGYKIDRIKGEIEFHNVTYFHPSPREVKILNLSMV 444
 QY 419 VOSGQTVAAVGNSSCGKSTTVOLMQLYDPDGMVCIQDGOIRITNVRHLEITGVVSOE 478
 Db 445 IKGETTAIFVSGSGAKSTALDILQRFDPCEGMYTLDGHIRSLNIRWLDQIGYIEOE 504
 QY 479 PVFATTIENIRYGENVTMDIEIKAVKANAYDFIMKLPNKPDITVIGRGAOLSGGOK 538
 Db 505 PVLFSTTIENIRLREEMEDIVQAAKDANAYNFIMALPQGDITLVGGGGOMSGGOK 564
 QY 539 ORIAIRALVRNPKILLDEATSAIDTESEAVNOVALDKARKGTTIVIAHRLSTVRND 598
 Db 565 ORVALARALIRKPKILLDMATSAIDNESEAKVQALNKIQHGHITIVSAHRLSTVRSAD 624
 QY 599 VTAGDDGYVAKGHHDELMKEKGIYFKLVYMTQGRNEIELENA--TGESSESDALEM-- 655
 Db 625 VTIQFEGHVAVERGHEHLEIRKKGIVFEMLVYLSQEDNTHKEITGIGKDTBEDTPERTF 684
 QY 656 ---SPKDSGSSILIKRSTR-----SIHAPQGDRLKTKEDILNENVPVSFW 700
 Db 685 SRGVSQDLSRLASIKRQSKSLSLSHSHEPPLAIDGHSYSYEDR--DNDAIVEVEVPAPVR 742
 QY 701 RLTKINSTEMPYVAVGFCALINGLOPAFSIIFSRILIGTFDEDEPETKRNQSMFVYL 760
 Db 743 RILKYINSEMPYIIVGALCAINGAVPIYSLFSGIILTFGL--VDKEQQRSEIYSKCLF 801
 QY 761 FLVLGIISFIFFLOGFTFGKAGEIILTKRLRYVFSMSLRQVSNFDDPKNTGATLTRL 820
 Db 802 FVILGCVSLFTQPLQGNFPAKSGEELTKRLRFGFAMLRQDIGNFDDLKNPGLVTRL 861
 QY 821 ANDAAQVKAIGSLAVITONIANLGTGIIISIIYGMQTLTLLAVPIIIMAGVEKEM 880
 Db 862 ATDAISOVQATGSGOVAMVNSFNIFFVAVLIAFLFNWKLISLVISFFFLALSAVQTKM 921
 QY 881 LSGQALKDKKELGAKIKATIEAENFRYVSLTRQKEVYVPAQSLQVYRNSLRKAHIF 940
 Db 922 LTGPAQDKKILEKAGQITINEALSNIRTVAGIGVGRIFKAFEELEKSYKTAIRKANVY 961
 QY 941 GVSPSTIQAAMVFSVAGCFRFGAVLVANEFMFOVLLVPSAIFYGMAVQVSSRPDY 1000
 Db 982 GLCYARSGISIFLANSAAVRGYLIVYEDINFSIVRVVSSILMSATAVAGRTFSYTPSY 1041

QY 1001 AAKYSAHVMIETKESPLDYSPPHGLKFNPLEGNVFNENYPTRPDIPVGLSL 1060
 DB 1042 AAKKISAAFFFLDRKRPDIIVYSGAEKWNFOGKIDICKFYPSPRPIQVNLST 1101
 QY 1061 EYKSGQTALVSSGCGKSTVQLLEFYDPLAGSVLIDGKEIKHLNOMLRAHLGVSQ 1120
 DB 1102 SYDPGQTALFVSSGCGKSTISQLLEFYDPOGTVIMIDGHSKKVNVQFLRSNIGIVSQ 1161
 QY 1121 EPLTPGCSIAENIANGDSRVYSHEEIMQAKAEANHHEFLPEKTYTRVGDGQSLG 1180
 DB 1162 EPLVDPGCSIMDKYIGDKETISVERALNAKQOLHDSVMSLPKETNVGIDGSQLSR 1221
 QY 1181 GOKORIALRALVROPHILLDEATSSALDTESEKVVQALDKAREGRTCIYIAHRLSTIQ 1240
 DB 1222 GOKORIALRALVROPHILLDEATSSALDTESEKVVQALDKAREGRTCIYIAHRLSTIQ 1281
 QY 1241 NADLIVFQNGKVEKHEGHOOILAOKGIFESWY 1273
 DB 1282 NSDIIAVMSQGVIEKGTGHRKINDQKAYYKLV 1314

RESULT 14

AB11_RAT STANDARD; PRT: 1321 AA.
 ID AB11_RAT
 AC 070127;
 DT 16-OCT-2001 (Rel. 40, Created)
 DT 16-OCT-2001 (Rel. 40, Last sequence update)
 DT 16-OCT-2001 (Rel. 40, Last annotation update)
 DE Bile salt export pump (Atp-binding cassette, sub-family B, member 11)
 GN ABCB11 OR BSEP OR SPCP
 OS Rattus norvegicus (Rat).
 OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.
 ON NCBI_TaxID=10116;
 RN (1)
 RP SEQUENCE FROM N.A.
 RC STRAIN=Sprague-Dawley; TISSUE=Liver;
 RX MEDLINE=98212048; PubMed=9545351;
 RA Gerloff T., Stieger B., Hagenbuch B., Madon J., Landmann L., Roth J.,
 RA Hofmann A.F., Meier P.J.;
 RA "The sister of P-glycoprotein represents the canalicular bile salt
 RT export pump of mammalian liver."
 RL J. Biol. Chem. 273:10046-10050(1998).
 CC -1- FUNCTION: INVOLVED IN THE ATP-DEPENDENT SECRETION OF BILE SALTS
 CC -1- INTO THE CANALICULUS OF HEPATOCYTES.
 CC -1- SUBCELLULAR LOCATION: Integral membrane protein.
 CC -1- TISSUE SPECIFICITY: EXPRESSED PREDOMINANTLY, IF NOT EXCLUSIVELY IN
 CC THE LIVER, WHERE IT WAS FURTHER LOCALIZED TO THE CANALICULAR
 CC MICROVILLI AND TO SUBCANALICULAR VESICLES OF THE HEPATOCYTES BY IN
 CC SITU.
 CC -1- DOMAIN: MULTIFUNCTIONAL POLYPEPTIDE WITH TWO HOMOLOGOUS HALVES,
 CC EACH CONTAINING AN HYDROPHOBIC MEMBRANE-ANCHORING DOMAIN AND AN
 CC ATP BINDING CASSETTE (ABC) DOMAIN.
 CC -1- SIMILARITY: BELONGS TO THE ABC TRANSPORTER FAMILY. MDR SUBFAMILY.
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC
 CC EMBL: U69487; AAC0084.1;
 DR InterPro: IPR003593; AAA_ATPase.
 DR InterPro: IPR003439; ABC_transport.
 DR InterPro: IPR001140; ABCtranspRTM.
 DR Pfam: PF00005; ABC_tran; 2.
 DR Pfam: PF00664; ABC_membrane; 2.
 DR ProDom: PD000006; ABC_transport; 2.
 DR SMART: SM00382; AAA; 2.

DR PROSITE: PS00211; ABC_TRANSPORTER; 1.
 KW ATP-binding; Transmembrane; Transport.
 FT DOMAIN 1 62
 FT TRANSMEM 63 83
 FT DOMAIN 84 147
 FT TRANSMEM 148 168
 FT DOMAIN 169 215
 FT TRANSMEM 216 236
 FT DOMAIN 237 240
 FT TRANSMEM 241 261
 FT DOMAIN 262 319
 FT TRANSMEM 320 340
 FT DOMAIN 341 353
 FT TRANSMEM 354 374
 FT DOMAIN 375 755
 FT TRANSMEM 756 776
 FT DOMAIN 777 794
 FT TRANSMEM 795 815
 FT DOMAIN 816 869
 FT TRANSMEM 870 890
 FT TRANSMEM 891 911
 FT DOMAIN 912 979
 FT TRANSMEM 980 1001
 FT DOMAIN 1001 1011
 FT TRANSMEM 1012 1032
 FT DOMAIN 1033 1321
 FT NP_BIND 455 462
 FT NP_BIND 1113 1120
 FT CAROHD 109 109
 FT CAROHD 116 116
 FT CAROHD 122 122
 FT CAROHD 125 125
 SQ SEQUENCE 1321 AA; 146257 MW; 543F4E7B9F9F1f6 CRC64;
 Query Match 49.5%; Score 3205.5; DB 1; Length 1321;
 Best Local Similarity 50.3%; Pred. No. 176-173;
 Matches 647; Conservative 233; Mismatches 372; Indels 35; Gaps 11;
 QY 20 KSKRNEK-KKKKPYSTFAMFRYSNMILRLTMLVGTMAALITHCAALPLMLVFGNMTDS 78
 DB 30 KSRIDQMKRDKIRGVGFELFRSSSKIKMLMGVCALLRHQMAQPGILIGINTDI 89
 QY 79 FANADISR-----KKTPEVLINES-----ITNTQHFHLEDEMTTAYVYSGI 123
 DB 90 FKTYIDENQLEITGKACVNNLT-VWINSFHOHNTGCTVGLVDISENINSGIYAGV 148
 QY 124 GAGVLVAAYIVGSFWSCLAAGROILTKRQFHAIKROEIGWFDVHDVDELNTRLDVSK 183
 DB 149 GMTVLILGFOIRLMVITGARQIRRMKTYFRRIIMKEIGWFDCTSVGELNSRFADIEK 208
 QY 184 INEGIGDKIGMFHSIATFFTGIVGFTGKMLTLYIAISPVLGSAIWAIRLISFTD 243
 DB 209 INDATADQLAHFIDQRMSTAMCGLLGFGYGRKLTVLIVLAVSPILGICAAVIGLSTKFT 268
 QY 244 KELLAAAGAGVAEYVLAIRYIARGOKRELYENKNELEAGIGIKRAITANISGA 303
 DB 269 LELKAYRAAGSIDAEVLSIRYAAFGENKKEVTEINELFAORWISIMGMGCEFTFY 328
 QY 304 AELLIVASALAEWGTSLV-SSEYSIGVLYTFEVSYLGAFSIGQASPIEAFANARG 362
 DB 329 MNCLEFFCYALAEWYGSITLDEEYTPGTVOIFLCVILAAIMNIGHASSCLEIFSTGCS 388
 QY 363 AAYEFKTIIDKPSIDBSYSGHPRNTRKNGLEPNVHVSFSPSKREKIKGLNKLVSQ 422
 DB 389 AATNIFQRIIDKQPIIDMGSDGKIDRKIGELFEHNATYFHPSPDKILDNLSMWIKNG 448
 QY 423 QTVALVNSGCGKSTYVQLMQRILDPIDGAVCIDGDIRITINVAHLREITGVVSQEVLE 482
 DB 449 ETTALVSSGSGKSTALQILQRFYDPCBEGWYLLDGHDIRSLINIMWLADQIGYDEPYLE 508
 QY 483 ATTIAENIRYGRNENYVDETERAVKEANAYDFIMKLPKKFTLIVGERGASQSLSGOKORIA 542
 DB 509 STTIAENIRFGEDATMEDIVAAKADANAYNFMALPQGFDTLVGEGSGMGSGQKORAVA 568

```

QY 543 IARALVNRKILLDEATSDALDTESEAVVOVALDKARKGRTTIVIAHRLSTVENADVING 602
Db 569 IARALIRNRKILLDMATISALDNESEARVOALNKHIGHITISVARIHSTVRAADVITIG 628
QY 603 FDDGVAVEKGNDELMEKRGYFKLVYMTQGRNEIELENA-TGESSESDALM-----S 656
Db 629 FEHGAVERGCHHEILLERKGYFMLVTLQSGDNAHETSIMKDKATEGGTLEKTFRSRG 688
QY 657 PDSGSSSLIKRSTR--SIHAP--QGDRKLGTKD-----LNENVPFVSFWRLTKLN 706
Db 689 YDLSIRASIRORSQSLSLTHDPPLAVADHKSYSKDSKNDVLYEVEPAPVRRLTKYN 748
QY 707 STEMYFVVGJFCALINGLOPAPSIIFSRIGITFREDPEPKRNSMFWLVVIGI 766
Db 749 IPEMHYILVGSLSAINGAVPIYSILFSQLGTFSL-LDKEQSEHSMCLFVVLIGC 807
QY 767 ISFTFFLGTFPGFKGILLTKRLRYVFRSMRLQOVSWFDDPKNTGALTTRLANDAQ 826
Db 808 VSIFLQGLQGTFAKSGSELLTKRLKRGFKAMLGODIGFDDLRRNPGVLTTRLATASQ 867
QY 827 VKGALGRVAVTQNIANLGTGIIISLIYGMOTLLIIVPIIAAGVVEKMLSGCAL 886
Db 868 VQGTGSGOVMMVWSFNITIAALLIFPFSWLSLITTFPPLALSGAVOTKMLTGPAS 927
QY 887 KDKKELGAGKIAETALENRTRYVSLTRQKFEYMAOSLOWPYRNSLRKAHIFGVSFSI 946
Db 928 QDNKALEKAGQITSEALSNIRYVAGVGRFKAFEEVELQTSKTAVRKANITIGLCFAF 987
QY 947 TQAMMYFVAGCFRGAVLYVANEFMFOVLLVFSAIYFGAMAGOVSSFPDYAKAVS 1006
Db 988 SSGIAFLANSAAVRYGGLIAYEGIGFSHVFRVVSVALSTAVGRFESSYTPSPAKAKIS 1047
QY 1007 AAHVIMIRKSPILDSYSHGKPKPTLGNTVNEVFNYPTRPDIVLQGLSLEVAKQ 1066
Db 1048 AAFEPULDRKPPILNYSVAGKMDNFOGKIDFIDCKFTYPSRBDIVLNGLSVAPGQ 1107
QY 1067 TALVSSGCGKSTVOLLERFYPDLASVGLDKETKHLNVQMLRAHLGVQOEPLTD 1126
Db 1108 TALVSSGCGKSTQGLERFYPDQGVIMIDHDKKYNIDPLRSNIGIYSQOEVLDR 1167
QY 1127 CSTAEINAVGDSRVVSHHEEIMQAKENIHFTETLPERKYNRVGDKGTGSGGOKORI 1186
Db 1168 CSTIMONIKGDMTKETIVERAIAAKOALHDFWMSLPEKYEINVGIOQSOSRGKOKOI 1227
QY 1187 AIARALVROPHILLDEATSDALDTESEKVVQVALDKAREBRTIVIAHRLSTQNALDV 1246
Db 1228 AIARALVROPHILLDEATSDALDTESEKTYQVALDKAREBRTIVIAHRLSTQNSDIIA 1287
QY 1247 VFQNGKVEKHTHOOLLAOKGITYFSWY 1273
Db 1288 VVSGGVTEKGTHERKMAOKGAYYKIV 1314

RESULT 15
MDRL CAEEL
ID MDRL CAEEL STANDARD; PRT: 1321 AA.
AC P34712;
DT 01-FEB-1994 (Rel. 28, Created)
DT 01-FEB-1994 (Rel. 28, Last sequence update)
DT 16-OCT-2001 (Rel. 40, Last annotation update)
DE Multidrug resistance protein 1 (P-glycoprotein A).
GN Pgp-1.
OS Caenorhabditis elegans.
OC Eukaryota; Metazoa; Nematoda; Chromadorea; Rhabditida; Rhabditidae;
OC Rhabditidae; Pelodierinae; Caenorhabditis.
OX NCBI_TaxID=6239;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=Bristol N2;
RA MEDLINE=93085750; PubMed=1360540;
RA Lincke C.R., The I., van Groenigen M., Borst P.;
RT "The P-glycoprotein gene family of Caenorhabditis elegans. Cloning

```

```

RT and characterization of genomic and complementary DNA sequences.";
RL J. Mol. Biol. 228:701-711(1992).
RN [2]
RP TISSUE SPECIFICITY.
RC STRAIN=Bristol N2;
RX MEDLINE=93223702; PubMed=8096815;
RA Lincke C.R., Brooks A., The I., Plasterk H.A., Borst P.;
RT "The expression of two P-glycoprotein (Pgp) genes in transgenic
RT Caenorhabditis elegans is confined to intestinal cells.";
RL EMBL J. 12:1615-1620(1993).
CC - FUNCTION: ENERGY-DEPENDENT EFFLUX PUMP RESPONSIBLE FOR DECREASED
CC - DRUG ACCUMULATION IN MULTIDRUG-RESISTANT CELLS.
CC - SUBCELLULAR LOCATION: Integral membrane protein.
CC - TISSUE SPECIFICITY: INTESTINAL CELLS.
CC - SIMILARITY: BELONGS TO THE ABC TRANSPORTER FAMILY, MDR SUBFAMILY.
CC This SWISS-PROT entry is copyrighted. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL: X65054; CAA46190.1; -.
DR PIR: S27337; S27337.
DR InterPro: IPR003593; AAA_ATPase.
DR InterPro: IPR003493; ABC_transport.
DR InterPro: IPR001140; ABCtransportTM.
DR Pfam: PF00005; ABC_tran; 2.
DR Pfam: PF00664; ABC_membrane; 2.
DR ProDom: PD000006; ABC_transport; 2.
DR SMART: SM00382; AAA; 2.
DR PROSITE: PS00211; ABC_TRANSPORTER; 2.
KW ATP-binding; Glycoprotein; Transmembrane; Transport; Repeat;
KW Multigene family.
FT DOMAIN 1
FT TRANSSEM 78 77 CYTOPLASMIC (POTENTIAL).
FT TRANSSEM 144 164 POTENTIAL.
FT TRANSSEM 213 233 POTENTIAL.
FT TRANSSEM 240 260 POTENTIAL.
FT TRANSSEM 321 341 POTENTIAL.
FT TRANSSEM 350 370 POTENTIAL.
FT DOMAIN 2
FT TRANSSEM 371 753 POTENTIAL.
FT TRANSSEM 754 774 CYTOPLASMIC (POTENTIAL).
FT TRANSSEM 798 818 POTENTIAL.
FT TRANSSEM 874 894 POTENTIAL.
FT TRANSSEM 895 915 POTENTIAL.
FT TRANSSEM 978 998 POTENTIAL.
FT TRANSSEM 1017 1037 POTENTIAL.
FT DOMAIN 1038 1321 POTENTIAL.
FT NP_BIND 451 458 CYTOPLASMIC (POTENTIAL).
FT NP_BIND 1112 1119 ATP (POTENTIAL).
FT NP_BIND 1321 1321 ATP (POTENTIAL).
SQ SEQUENCE 1321 AA; 145074 MW; 6130AAFB3B1A8FA9 CRC64;

Query Match 42.3%; Score 2739; DB 1; Length 1321;
Best Local Similarity 45.0%; Pred. No. 4, 1e-147;
Matches 584; Conservative 221; Mismatches 434; Indels 58; Gaps 12;

QY 21 KSKRKEKKEKPYSTRAMRYSNWDLRYMTAIIHGALPLMLVPGNNTDSFA 80
Db 47 KITRKAEEVNVKVSIPOLVRYTTLEKLLFGLVAVTGGAGLWMSILOGKRSQAFI 106
QY 81 NAGISRNKTFPVINISITNNTOHFI-----NLLEEMTYAYVYSGAGVLYVAY 132
Db 107 NE-----QIVIN-----NNGSTFLPTGQNTKTDFEDVNVVSYAAMTVGMMAGQ 154
QY 133 IOVSFWCLAQRQILKIRKOFPAIRKQETIGWFDVHVGELNTRLTVDYSKINGIGDKI 192
Db 155 IIVCYLYVAEQNNNRNIRREFVKSILROEISWFDNHSGLTATKFLDNLKRYKGTGDKI 214
QY 193 GMEFHSIAFFETFIYGTGKWLTLVILAIISPVLGSAALWAKILSSFDKELAYAKA 252

```


GenCore version 5.1.3
Copyright (c) 1993 - 2002 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: December 9, 2002, 16:20:53 ; Search time 1313 Seconds

(without alignments)
201.025 Million cell updates/sec

Title: US-09-672-725C-2

Perfect score: 6477
Sequence: 1 MDPEGRKRSASAKNFWKMGK.....LLAQKGIYFSWVSQAGAKR 1281

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 671580 segs, 206047115 residues

Total number of hits satisfying chosen parameters: 671580

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Database : Listing first 45 summaries

1: SP_ARCHAEA:*
2: SP_BACTERIA:*
3: SP_FUNGI:*
4: SP_HUMAN:*
5: SP_INVERTEBRATE:*
6: SP_MAMMAL:*
7: SP_MHC:*
8: SP_ORGANELLE:*
9: SP_PHAGE:*
10: SP_PLANT:*
11: SP_RODENT:*
12: SP_VIRUS:*
13: SP_VERTEBRATE:*
14: SP_UNCLASSIFIED:*
15: SP_VIRUS:*
16: SP_BACTERIAP:*
17: SP_ARCHAEP:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	6420.5	99.1	1280	6	046605
2	5682.5	87.7	1285	6	002793
3	5629	86.9	1272	11	09JUK64
4	5607	86.6	1272	11	0924L6
5	5516.5	85.2	1163	6	09T5U2
6	5333	82.3	1169	11	060502
7	5286.5	81.6	1275	11	08R4Z7
8	4660.5	72.0	1288	13	093437
9	4430	68.4	1287	13	091586
10	3331	51.4	1348	13	091386
11	3315	51.2	668	6	08WNU6
12	3209.5	49.6	1321	11	088331
13	2857	44.1	1294	5	045721
14	2852.5	44.0	1407	5	08T9W5
15	2827	43.6	1265	5	001495
16	2825	43.6	851	13	09W693

17	2818.5	43.5	1275	5	061301
18	2751.5	42.5	1432	5	08T9W4
19	2741	42.3	1321	5	021349
20	2701	41.7	1313	5	09V6Z6
21	2652	40.9	1283	5	024393
22	2612.5	40.3	817	13	090WV6
23	2603.5	40.2	1285	10	09LGX1
24	2602.5	40.2	1286	10	080725
25	2596	40.1	1292	10	09M109
26	2589.5	40.0	1278	10	09FWX7
27	2557	39.5	1279	5	09V616
28	2540.5	39.2	1262	10	08RV77
29	2532	39.1	1348	3	09Y8G1
30	2521	38.9	1229	10	09SVY2
31	2518	38.9	1229	10	049749
32	2518	38.9	1230	10	09SVY3
33	2517	38.9	1289	10	094TH6
34	2511	38.8	1349	3	043121
35	2489.5	38.4	1287	10	08S035
36	2473.5	38.2	1313	10	09ZRG2
37	2473	38.2	1229	10	09FWX8
38	2465	38.1	1292	3	09C163
39	2441	37.7	1408	3	043140
40	2437.5	37.6	1323	10	09W0M2
41	2431	37.5	1286	10	09ZKR7
42	2423	37.4	1310	5	024852
43	2419.5	37.4	1248	10	09FHF1
44	2418.5	37.3	1331	3	09HGF5
45	2410	37.2	1252	10	09LTX0

ALIGNMENTS

RESULT 1

046605 PRELIMINARY; PRT; 1280 AA.
ID 046605:
AC 01-JUN-1998 (TREMBLER, 06, Created)
DT 01-JUN-1998 (TREMBLER, 06, Last sequence update)
DT 01-JUN-2002 (TREMBLER, 21, Last annotation update)
DE Multidrug resistance P-glycoprotein.
GN MDRL.
OS Canis familiaris (Dog).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Carnivora; Fissipedia; Canidae; Canis.
OX NCBI_TaxID=9615;
RN [1]
RP SEQUENCE FROM N.A.
RA TISSUE=COLON;
RC Pucl O.; Lepage J.F., Alvine M., Galtier P., Pineau T.;
RT Submitted (JAN-1998) to the EMBL/GenBank/DBJ databases.
CC -1- SIMILARITY: BELONGS TO THE ABC TRANSPORTER FAMILY.
DR EMBL: AF045016; AAC02113.1; -
DR InterPro: IPR003593; AAA_Arpase.
DR InterPro: IPR003439; ABC_transportr.
DR Pfam: PF00664; ABC_membrane_2.
DR Pfam: PF00005; ABC_tran; 2.
DR ProDom: PD000006; ABC_transportr; 2.
DR SMART: SM00382; AAA; 2.
DR PROSITE: PS00211; ABC_TRANSPORTER; 2.
DR ATP-binding; Transport.
KW SEQUENCE 1280 AA; 141524 MW; 762DD5AF4C73306 CRC64;

Query Match 99.1%; Score 6420.5; DB 6; Length 1280;

Best Local Similarity 99.3%; Pred. No. 0;

Matches 1272; Conservative 4; Mismatches 4; Indels 1; Gaps 1;
OY 1 MDPEGRKRSASAKNFWKMGKSKK...EKKKKPTVSTFAMFRYSNWLDRLYLVTMAAIT 60
DB 1 MDPEGRKRSASAKNFWKMGKSKK...EKKKKPTVSTFAMFRYSNWLDRLYLVTMAAIT 59

61 HGAALPLMLVFGNMTDSFANAGISRNKTPVLIINESITNNQHFINHLEEMTYAAYX 120
DB 60 HGAALPLMLVFGNMTDSFANAGISRNKTPVLIINESITNNQHFINHLEEMTYAAYX 119
QY 121 SGIGAGVVAAYIOVSFMCIAAGROIILKIRKOFPHAIMROEIGWFDVHDVDELNRLDJD 180
DB 120 SGIGAGVVAAYIOVSFMCIAAGROIILKIRKOFPHAIMROEIGWFDVHDVDELNRLDJD 179
QY 181 VSKINEGIGDKIGMEFHSIATFETGFIYGFTRGKLLVLIATISPVIGLSAAIIMAKILSS 240
DB 180 VSKINEGIGDKIGMEFHSIATFETGFIYGFTRGKLLVLIATISPVIGLSAAIIMAKILSS 239
QY 241 FPDKELLAAKAGAAVEVLAIRTVIAFGOKKELERYNKNLEAKGIGIKKAITANIS 300
DB 240 FPDKELLAAKAGAAVEVLAIRTVIAFGOKKELERYNKNLEAKGIGIKKAITANIS 299
QY 301 IGAAPLLIYASALAFWYGTSLVSESYSIQOVLTFPFSVLIGAFSIGQASPSIEARANA 360
DB 300 IGAAPLLIYASALAFWYGTSLVSESYSIQOVLTFPFSVLIGAFSIGQASPSIEARANA 359
QY 361 RGAAYIEFKIIDNKPSIDSYSGSKPDKIKGNLEFKVNHSTYSRKEVKILKGLNLKVQ 420
DB 360 RGAAYIEFKIIDNKPSIDSYSGSKPDKIKGNLEFKVNHSTYSRKEVKILKGLNLKVQ 419
QY 421 SGGTVALVNGSGCKSTVOLMORLYDPTDQWVCIDGODIRFTINVRHLREITGVYSGEPV 480
DB 420 SGGTVALVNGSGCKSTVOLMORLYDPTDQWVCIDGODIRFTINVRHLREITGVYSGEPV 479
QY 481 LEATTIAENIRIGRENVYDELEKAVKEMANADFTIMKIPNKFDLYVERGALSGGOKOR 540
DB 480 LEATTIAENIRIGRENVYDELEKAVKEMANADFTIMKIPNKFDLYVERGALSGGOKOR 539
QY 541 IATARALVRNPKILLDEATSAIDTESAVVOALDKARKGRTTIVIAHRLSTVANADVI 600
DB 540 IATARALVRNPKILLDEATSAIDTESAVVOALDKARKGRTTIVIAHRLSTVANADVI 599
QY 601 AGEDDGVIVEKGNHDELMKEKGIYFLVLMQTRNGEIELENNATGSKESDLESPDS 660
DB 600 AGEDDGVIVEKGNHDELMKEKGIYFLVLMQTRNGEIELENNATGSKESDLESPDS 659
QY 661 GSSLIKRSTRRSIIHAPOGODRKLGTEDINENVPVSFWIRILKINSTEMPFVVGIFCA 720
DB 660 GSSLIKRSTRRSIIHAPOGODRKLGTEDINENVPVSFWIRILKINSTEMPFVVGIFCA 719
QY 721 IINGLOPAFSIIFSRILIGITRDEDEPFRKONSNEVFLVYLGLISTITFELIGFTFG 780
DB 720 IINGLOPAFSIIFSRILIGITRDEDEPFRKONSNEVFLVYLGLISTITFELIGFTFG 779
QY 781 KAGELTRKRLRYWYERSMRLRODYSWFPDDPKNTGALTTRLANDAQAIGSRILAVIQ 840
DB 780 KAGELTRKRLRYWYERSMRLRODYSWFPDDPKNTGALTTRLANDAQAIGSRILAVIQ 839
QY 841 NIANGTIIISLIYGMQLTLLLAIVPIAIAGVVEMKMLSGOALKDKKLEGAKIAT 900
DB 840 NIANGTIIISLIYGMQLTLLLAIVPIAIAGVVEMKMLSGOALKDKKLEGAKIAT 899
QY 901 EAIENFRVVSILTRBOKFEYMAOSLOVPTRNSLRKAHIFGSEFSTIOAMMYSTVACGR 960
DB 900 EAIENFRVVSILTRBOKFEYMAOSLOVPTRNSLRKAHIFGSEFSTIOAMMYSTVACGR 959
QY 961 FGAYIVANEFNFDVLIVFSAIYFGAMAVGOVSSFAADYAKAKVSAHAHYIMITEKSPIL 1020
DB 960 FGAYIVANEFNFDVLIVFSAIYFGAMAVGOVSSFAADYAKAKVSAHAHYIMITEKSPIL 1019
QY 1021 DSYSPHGLKPTLLEGNVTTFNEVYNYPTRPDIPIVLOGLSLIEVKKGGOLAVVSGGCKST 1080
DB 1020 DSYSPHGLKPTLLEGNVTTFNEVYNYPTRPDIPIVLOGLSLIEVKKGGOLAVVSGGCKST 1079
QY 1081 VVOLLERYPDLPLAGSVLIDGKEIKHLNWOHLRAHGIIVSOEPIIFDCSIAENIAYGDNRSR 1140
DB 1080 VVOLLERYPDLPLAGSVLIDGKEIKHLNWOHLRAHGIIVSOEPIIFDCSIAENIAYGDNRSR 1139
QY 1141 VVSHEEIMQAKKANIHHFIETLPEKITYTRVGDXTOLSGGOKORIALAIALVRQPHILL 1200

DB 1140 VVSHEEIMQAKKANIHHFIETLPEKITYTRVGDXTOLSGGOKORIALAIALVRQPHILL 1199
QY 1201 IDEATSAIDTESKVOEALDKAREGRTICIVIAHRLSTIONADLIYVPONGKVEKHTHQ 1260
DB 1200 IDEATSAIDTESKVOEALDKAREGRTICIVIAHRLSTIONADLIYVPONGKVEKHTHQ 1259
QY 1261 QLLAQKGIYFSMVSOAGAKR 1281
DB 1260 QLLAQKGIYFSMVSOAGAKR 1280
RESULT 2
002793 PRELIMINARY; PRT; 1285 AA.
ID 002793
AC 002793
DT 01-JUL-1997 (TREMBLrel. 04, Created)
DT 01-JUL-1997 (TREMBLrel. 04, Last sequence update)
DT 01-JUN-2002 (TREMBLrel. 21, Last annotation update)
DE Multidrug resistance protein-1.
GN MDRL.
OS Ovis aries (Sheep).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Cetartiodactyla; Ruminantia; Pecora; Bovidae;
OC Bovidae; Caprinae; Ovis.
OX NCBI_Taxid=9940;
RN [1]
RP SEQUENCE FROM N.A.
RA Longley M., Crawford A.M.:
RT "Ovine mdrl gene."
RL Submitted (NOV-1996) to the EMBL/GenBank/DBJ databases.
RC -1- SIMILARITY: BELONGS TO THE ABC TRANSPORTER FAMILY.
DR EMBL: U78609; AAB58489.1; -
DR InterPro: IPR003593; AAA_Arpase.
DR InterPro: IPR001140; ABCtransport.
DR InterPro: IPR003439; ABC_transport.
DR Pfam: PF00664; ABC_membrane_2.
DR Pfam: PF00005; ABC_tran_2.
DR ProDom: PD000006; ABC_transportr; 2.
DR SMART: SM00382; AAA; 2.
DR PROSITE: PS00211; ABC_TRANSPORTER; 2.
KW ATP-binding; Transport.
SQ SQUONCE 1285 AA; 142020 MW; 90153617C44856F CRC64;
Query Match 87.7%; Score 5682.5; DB 6; Length 1285;
Best Local Similarity 87.3%; Pred. No. 0;
Matches 1122; Conservative 80; Mismatches 78; Indels 5; Gaps 3;
QY 1 MDPEGRKKG-SARKNMKMGK---SKNNEKKKRPVSTFAMFRYSNMLDRLVLMVGTM 56
DB 1 MDLEGDRNGRAAGNLFKRDKKRFFSKDKKKEKRPTVSTFTMFRRYSNMLDRLCYVLGTL 60
QY 57 AAIHGAALPLMLVFGNMTDSFANAGISRNKTPVLIINESITNNQHFINHLEEMTY 116
DB 61 AAIHGAALPLMLVFGDMTDSFAGAGNLGNTILSNISTSTIDRTE-YGKLEKEMTY 119
QY 117 AAYTSGAGVVAAYIOVSFMCIAAGROIILKIRKOFPHAIMROEIGWFDVHDVDELNRL 176
DB 120 AAYTSGAGVVAAYIOVSFMCIAAGROIILKIRKOFPHAIMROEIGWFDVHDVDELNRL 179
QY 177 LTDDVSKINEGIGDKIGMEFHSIATFETGFIYGFTRGKLLVLIATISPVIGLSAAIIMAK 236
DB 180 LTNDVSKINEGIGDKIGMEFHSIATFETGFIYGFTRGKLLVLIATISPVIGLSAAIIMAK 239
QY 237 ILSFTDKELLAAKAGAAVEVLAIRTVIAFGOKKELERYNKNLEAKGIGIKKAIT 296
DB 240 ILSFTDKELLAAKAGAAVEVLAIRTVIAFGOKKELERYNKNLEAKKRRIGIKKAIT 299
QY 297 ANISIGAAPLLIYASALAFWYGTSLVSESYSIQOVLTFPFSVLIGAFSIGQASPSIEA 356
DB 300 ANISIGAAPLLIYASALAFWYGTSLVSESYSIQOVLTFPFSVLIGAFSIGQASPSIEA 359
QY 357 FANAGAAVEIEFKIIDNKPSIDSYSGSKHPDKIKGNLEFKVNHSTYSRKEVKILKGLNL 416

Dd	360	FANMAGAAVEFKKILLNDKPSIDSYNTWGHKPNKNGNLEFRVNEHHPYSRENEVKKLGLN	419
Qy	417	LKVSGQGTVALVNGSGCGKSTTVQLMQRLYDPTDQAVCIDGODIRTNVRLREITGVYS	476
Dd	420	LKVSGCGTVALVNGSGCGKSTTVQLMQRLYDPTDQAVCIDGODIRTNVRLREITGVYS	479
Qy	477	QEPVLFATTAENIRRGRENVTMDETEKAVKEANAIDEIMKLPNKFTLVGERGAQLSGG	536
Dd	480	QEPVLFATTAENIRRGREDVYTMDEIQKAVKEANAIDEIMKLPNKFTLVGERGAQLSGG	539
Qy	537	QKQRIATARALVNRPKILLIDETSALDTESEAAVVOVALDKARKGRFTTYIAHRLSTVRN	596
Dd	540	QKQRIATARALVNRPKILLIDETSALDTESEAAVVOVALDKARKGRFTTYIAHRLSTVRN	599
Qy	557	ADVIAGDDGVIVEKNHDELKENGKGYIFKVLVTMQTRNEITEENATGEKSSSDLEMS	656
Dd	600	ADVIAGDDGVIVEEGSHDELKMGKGYIFKVLVTMQTRNEITEENATGEKSSSDLEMS	659
Qy	657	PKDGSLLIKRRSTRSRSHAPGQGRKLTGEDLNENVPVSEWRILKLNSTEMPVYVG	716
Dd	660	SQDSRSLLIKRRSTRSRSHAPGQGRKLTGEDLNENVPVSEWRILKLNSTEMPVYVG	719
Qy	717	IFCAILNGLOPAFSIIFSRIGIPTRDEDETEKRONSNFSLFVLITGLISTFTFFLOG	776
Dd	720	VFCAILNGLOPAFSVTFSRIGIPTRDNDETEKRONSNFSLFVLITGLISTFTFFLOG	779
Qy	777	FTFGKAGELLTKRLRMFPRSKLRQDVSFDPDKNTTGALTTRLANDAAQVGAIGSRLLA	836
Dd	780	FTFGKAGELLTKRLRLVFRSMRLRDVSWFDPDKNTTGALTTRLANDAAQVGAIGSRLLA	839
Qy	837	VITQNLANTGTIIISLIYGMQLTLLIATVPIIAINGVEMAMLSGQALKKCKELEAG	896
Dd	840	VITQNLANTGTIIISLIYGMQLTLLIATVPIIAAGVTEMMKLSGQALKKCKELEAG	899
Qy	897	KATKATEIENFRVVSILTRGEQFEYMAQSLQVPRNSLRAKHAHFGVSFIITQAMMYESEA	956
Dd	900	KATEIENFRVVSILTRGEQFEYMAQSLQVPRNSLRAHFGVTSTITQAMMYESEA	959
Qy	957	GCFRFAVLYVAEENFDVLLVSALVFGAMAVGVSSFPADPYAKAKVSAAHVTIMTEK	1016
Dd	960	GCFRGAIVLYVAOIMEFDVLLVSAVFAMAVGVSSFPADPYAKAKVSAAHVTIMTEK	1019
Qy	1017	SPILIDSYSPHGLKPTLEGNATFEVEVFNPTRPDIPLVGLSILEYKKGQTLALVSSGC	1076
Dd	1020	IFLIDSYSTEGKIDSTVGSVAIFADVFNPTRPDPVPLRGLSLEKKKGQTLALVSSGC	1079
Qy	1077	GKSTVQVLLERFPDPLAGSVLIDGKEIKHLNVQMLRAHLGIYSQBPILFDCSIAMENIAYG	1136
Dd	1080	GKSTVQVLLERFPDPLAGVYFIDGKEIKVQLNVQMLRAHMGIVSQBPILFDCSIAMENIAYG	1139
Qy	1137	DNSRVVSHHEEIMOAKEANIHHEFLLPEKYNVRVGDCKTQSLSGQKORIAIARALVROP	1196
Dd	1140	DNSRVVSEELHAKKANHSHFTFEMLPKYNVRVGDCKTQSLSGQKORIAIARALVROP	1199
Qy	1197	HILLDEATSLDTESEKVOEALDKARSGRTCIYIARHSLSTIONADLIYFONKGYKEH	1256
Dd	1200	HILLDEATSLDTESEKVOEALDKARSGRTCIYIARHSLSTIONADLIYFONKGYKEH	1259
Qy	1257	GTHOOLLAOGGIYFSNYSVQAGAKR 1261	
Dd	1260	GTHOOLLAOGGIYFTMVSQVAGTKR 1284	
RESULT 3			
O9JK64			
ID	O9JK64	PRELIMINARY:	PRT, 1272 AA.
AC	O9JK64		
DT	01-OCT-2000 (Tremblrel, 15, Created)		
DT	01-OCT-2000 (Tremblrel, 15, Last sequence update)		
DT	01-JUN-2002 (Tremblrel, 21, Last sequence update)		
DE	Multidrug resistance protein 1a.		
EN	PGY1.		

OS *Rattus norvegicus* (Rat).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus
OX NCBI_TaxID=10116;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=H1STAR; TISSUE=LIVER;
RA Hooiveld G.J.E.J., Wilms J.W.J., Hagenbuch B., Jansen P.L.M.,
RA Meijer D.K.F., Muller M.;
RT "Cloning and functional characterization of the rat multidrug
RT resistance protein Mdr1a."
RL Submitted (Apr-2000) to the EMBL/GenBank/DBJ databases.
CC -1- SIMILARITY: BELONGS TO THE ABC TRANSPORTER FAMILY.
DR EMBL, AF257746; AAF69007.1; -
DR InterPro: IPR003593; AAA_ATPase.
DR InterPro: IPR001140; ABCtransp1TM.
DR InterPro: IPR003439; ABC_transp1r.
DR Pfam: PF00664; ABC_membrane; 2.
DR Pfam: PF00005; ABC_tran; 2.
DR Prodom: PD00006; ABC_transp1r; 2.
DR SMART: SM00382; AAA; 2.
DR PROSITE, PS00211; ABC_TRANSPORTER; 2.
KW ATP-binding; Transport.
SQ SEQUENCE 1272 AA; 140326 MW; B5C1DE98FC986C5 CRC64.

Query Match	86.6%, Score 5607, DB 117, Length 1272:
Best Local Similarity	86.1%, Pred. No. 9, 2e-317:
Matches 1103; Conservative	80; Mismatches 88; Indels 10; Gaps
QY	1 MDPEGKRGKSAEKNFMKMGKSKKNEKKEKKPPYSTFAFRFRSMILDRILYMLVGTMAII 60
DB	1 MELEEDINGADNNSKMGKSKK EKKEKKRPAVSVLTMRVAGMLDRFYMLGLTIAII 59
QY	61 HGALPLMLMLVEGNMDSFPANAGISRNKTPVLIINESITNNIOHFNLHLEEMTYAAY 120
DB	60 HGIALPLMLMLVFGDMTDSFPANGNRSTSF-----YNATDIYARLEEDMTYAIY 110
QY	121 SGICAGVLAAYTVQSFNCLAGROILIKRQFHALMROEIMFVDHVGELNRLND 180
DB	111 TGICAGVLAAYTVQSLWCLAGROIHKROKFFHALMNOEIMFVDHVGELNRLND 170
QY	181 VSKINEGIGDKIMFHSIATEFTFTFVIGTRGKMLTVLILASPVLGISAIAIAKIISS 240
DB	171 VSKINEGIGDKIMFQOAMTEFGGPIIGTRGKMLTVLILASPVLGISAGIAKIISS 230
QY	241 FTDEKLLAVAKAGVAEEVLAIRVIYAFGCKKKELEYNKNLEEKAGIGIKKAIYANIS 300
DB	231 FTDEKLLQYAAAGVAEEVLAIRVIYIAGGKKKELEYNKNLEEKAKRILGIIKAIYANIS 290
QY	301 IGAALFLIYASIALAFMYGSIIVLSSEKSIIGOVLTVEFSYVIGAFSIGQASPIEAFANA 360
DB	291 MGAALFLIYASYIALAFMYGSIIVLSSEKSIIGOVLTVEFSYVIGAFSIGQASPIEAFANA 350
QY	361 RGAAYEIFKIIDNKPISDSTISKSGHPRDNIKNLEFKVHVSYSREKVEIILGLMLKY 420
DB	351 RGAAYEVSIIIDNKPISDSTISKSGHPRDNIOGNLEFKNIHVSYSRKVDYIILGLMLKY 410
QY	421 SGQYVALVNGSGCKSTYTVQLMORLDPDNGWCIDGODIRTIIVRHLREITGVSOEY 480
DB	411 SGQYVALVNGSGCKSTYTVQLLORLDPLEGEVSIIDGODIRTIIVRHLREITGVSOEY 470
QY	481 LEFATTIAENIRYGBENVTMDIEKAYKANAYDITMKLPNKPFDPLVGBRGNOISGGQOR 540
DB	471 LEFATTIAENIRYGBENVTMDIEKAYKANAYDITMKLPNKPFDPLVGBRGNOISGGQOR 530
QY	541 IAAARLAVNPKIILLBEATDSEEAVALDDEAKARKGRTTVIJAHLSTVRNADY 600
DB	531 IAAARLAVNPKIILLBEATDSEEAVALDDEAKARKGRTTVIJAHLSTVRNADY 590
QY	601 AGEFDDGVIEKGNHDELMKEKGYFKLVITMOTRGNFTELEANAIGSKSKESDALPMSKOS 660
DB	591 AGEFDDGVIEKGNHDELMKEKGYFKLVIAOTRGNFTELEANAIGSKSKESDALPMSKOS 650
QY	661 GSSLIKRRSTRRSIHAPGODRKIGTDEENENAVPVSFMRILIKNSTEMPEFYVVGIFCA 720
DB	651 GSSLIKRRSTRRSIHAPDQGEISTEALDDVPAASHMRILIKNSTEMPEFYVVGIFCA 710
QY	721 IINGGLOPAFSITFSIRIGITRDEDEPFRKONSNMFVYLVLGIIISFTIFFIQGTFFG 780
DB	711 IINGGLOPAFSITFSKVVGVFTKNDPDEIORONSNIISLFTLIIISFTIFFIQGTFFG 770
QY	781 KAGEIILTKRLRYVRSMRLRODVSEMPDPRKNTGATLTRLANDAAQVGAIGSRILAVTQ 840
DB	771 KAGEIILTKRLRYVRKSMRLRODISFPDPKRTIGALTTRLANDAAQVGAIGSRILAVTQ 830
QY	841 NIAMLGIIISITLYGMQTLTLLILAVIPIIAIAGVEMKMLSGOALKDKKELEGSGKAT 900
DB	831 NIAMLGIIISITLYGMQTLTLLILAVIPIIAIAGVEMKMLSGOALKDKKELEGSGKAT 890
QY	901 EAIENFTTVSLTRQEKFEYMTAQSLQVPPYNSLRKAHIGVSEFSTQAMVFYSVAGFER 960
DB	891 EAIENFTTVSLTRQEKFEYMTAQSLQVPPYNSLRKAHIGVIFSTQAMVFYSVAGFER 950
QY	961 FGALVLVNEEMODVLVFSALIVGAAAVQVSSFPDYAKAKASAAHIMIEKSPLI 1020
DB	951 FGALVLVRELMTEENVLVFSALIVGAAAVQVSSFPDYAKAKASASHIRIIEKPEI 1010
QY	1021 DSYSPHKLKPNLTLEGNVTNEVFNVPTRPDI PVYLGSLSEVKKGGOTLALVSSGCGKST 1080


```
Db 1011 DSYTBGKJPMLEGNKFNKGVMEFPTREPNIPVLOGLSLEVKGGOTLALVSSCGCKST 1070
Qy 1081 VVOLLERFYDPLAGSVLIDGKEIKHLNVOMLRAHLGIVSOEPIIDPCSAENIAAGDMSR 1140
Db 1071 VVOLLERFYDPMAGVTELDCKEIKOLNVOMLRAHLGIVSOEPIIDPCSAENIAAGDMSR 1130
Qy 1141 VVSHHEELMOAKKAPANIHHFETLPEKYNTRVGDKGTQLSGGOKORAIARALVROPHTLL 1200
Db 1131 VVSHHEELVRAKAKENIHOFIDSLPEKYNTRVGDKGTQLSGGOKORAIARALVROPHTLL 1190
Qy 1201 LDEATSAIDTESKVVQOALDKARBSERTCIYIAHRLSTIQANDLIVFPQNGKVEHGTQ 1260
Db 1191 LDEATSAIDTESKVVQOALDKARBSERTCIYIAHRLSTIQANDLIVFPQNGKVEHGTQ 1250
Qy 1261 QLIAQKGIYFSMVSVOAGAKR 1281
Db 1251 PLIAQKGIYFSMVSVOAGAKR 1271
```

RESULT 5

```
Q9TSU2 PRELIMINARY: PRT: 1163 AA.
AC Q9TSU2;
DT 01-MAY-2000 (Tremblrel. 13, Created)
DT 01-MAY-2000 (Tremblrel. 13, Last sequence update)
DT 01-JUN-2002 (Tremblrel. 21, Last annotation update)
DE Multi-drug resistance related (Fragment).
OS Felis silvestris catus (Cat).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Carnivora; Fissipedia; Felidae; Felis.
OX NCBI_Taxid=9685;
RN [1]
RP SEQUENCE FROM N.A.
RA Okada Y., Nakamura N., Matsusiro H., Kato H., Setoguchi A., Yazawa M.,
RA Okada M., Wataai T., Hasegawa A., Tsujimoto H.;
RT "Molecular analysis of multidrug resistance in lymphoma cells in the
RT cat.";
RT
CC Am. J. Vet. Res. 0:0-0(1999).
CC -1 SIMILARITY: BELONGS TO THE ABC TRANSPORTER FAMILY.
DR EMBL: AB029153; BA87071.1; -.
DR InterPro: IPR003593; AAA_ATPase.
DR InterPro: IPR001140; ABCtransportTM.
DR Pfam: PF00664; ABC_membrane_2.
DR Pfam: PF00005; ABC_tran_2.
DR ProDom: PD000006; ABC_transport_2.
DR SMART: SM00382; AAA_1.
DR PROSITE: PS00211; ABC_TRANSPORTER; 2.
KW ATP-binding; Transport.
FT NON_TER 1
FT TER 1163
SQ SEQUENCE 1163 AA; 128511 MW; 1B5B413776A93A26 CRC64;
```

Query Match 85.2%; Score 5516.5; DB 6; Length 1163;
Best Local Similarity 94.2%; Pred. No. 1.4e-311;
Matches 1096; Conservative 34; Mismatches 33; Indels 1; Gaps 1;

```
Qy 61 HGALPLMLMVFEGNMTDSFNAGISRNKTPPYVINSITNNQHPINLHEEMTYAYYY 120
Db 1 HGALPLMLMVFEGNMTDSFNAGISRNKTPPYVINSITNNQHPINLHEEMTYAYYY 59
Qy 121 SGIGAGVLAAYVIOVSFWCLAGRQILKTRKQFHAIMROEIGMFDVHDVGLNTRLDD 180
Db 60 SGIGAGVLAAYVIOVSFWCLAGRQILKTRKQFHAIMROEIGMFDVHDVGLNTRLDD 119
Qy 181 VSKINGIDGKTCMFHSHATFTFTGFIYFTGKMLFVILIAISVGLISAIAWAKILSS 240
Db 120 VSKINGIDGKTCMFHSHATFTFTGFIYFTGKMLFVILIAISVGLISAIAWAKILSS 179
Qy 241 FTDKELLAAVAKGAAVEEYLAIRTVIAFGGOKKLELYNNLLEAKIGIKKAITANIS 300
Db 180 FTDKELLAAVAKGAAVEEYLAIRTVIAFGGOKKLELYNNLLEAKIGIKKAITANIS 239
```

```
Qy 301 IGAFFLLIYASVALAWGTSTIVLSSEYSIGOVLTWFSVILIGAFSGOASPIAPANA 360
Db 240 IGIATFLIYASVALAWGTSTIVLSSEYSIGOVLTWFSVILIGAFSGOASPIAPANA 299
Qy 361 RGAAYEFKTIIDKPSIDYSKSGHAPDVIKGNLEKKNVHESYPSRKEVKKILKGLNKVQ 420
Db 300 RGAAYEFKTIIDKPSIDYSKSGHAPDVIKGNLEKKNVHESYPSRKEVKKILKGLNKVQ 359
Qy 421 SGQVALVNSGGCKSTVQLOMRLYDPTDGAWCIDGODIRITINVRHILREITGVSOEPP 480
Db 360 SGQVALVNSGGCKSTVQLOMRLYDPTDGAWCIDGODIRITINVRHILREITGVSOEPP 419
Qy 481 LEFTTIAENIRYGRENVYMDIEKAVKANAYDTMKLPNKFDTLVGRGAOLSGGOKOR 540
Db 420 LEFTTIAENIRYGRENVYMDIEKAVKANAYDTMKLPNKFDTLVGRGAOLSGGOKOR 479
Qy 541 IATARALVNRPKILLDEATSAIDTESKVVQOALDKARBSERTCIYIAHRLSTIRNADVI 600
Db 480 IATARALVNRPKILLDEATSAIDTESKVVQOALDKARBSERTCIYIAHRLSTIRNADVI 539
Qy 601 AGFDGVIVKGNHDELMKEKGIYKLTVMOTRGNEIELENAATESKSESDALEMSPKDS 660
Db 540 AGFDGVIVKGNHDELMKEKGIYKLTVMOTRGNEIELENAATESKSESDALEMSPKDS 599
Qy 661 GSSLKRSTRSIHAPGODRKLGCTKEDLNEVNPVPSFWRIKLNSTEMPYFVYGIFCA 720
Db 600 GSSLKRSTRSIHASOGODRKLGCTKENLEDPVPSFWRIKLNTITEMPYFVYGIFCA 659
Qy 721 IINGLOPAFSTIFSRIGITRDEDPETKRONSMFSVLEIVGITSITTFPLOGTFTG 780
Db 660 IINGLOPAFSTIFSRIGITRDEDPETKRONSMFSVLEIVGITSITTFPLOGTFTG 719
Qy 781 KAGELLTRKLRNVPFRSMLRODVSFDDPKNTGALTTRLANDAQAQKAGISRLAVITQ 840
Db 720 KAGELLTRKLRNVPFRSMLRODVSFDDPKNTGALTTRLANDAQAQKAGISRLAVITQ 779
Qy 841 NIANGTGIISLIYGMQTLTLLAIVPIITAIAGVEMKMLSQALDKKKELGACKIAT 900
Db 780 NIANGTGIISLIYGMQTLTLLAIVPIITAIAGVEMKMLSQALDKKKELGACKIAT 839
Qy 901 EAIENFTVSVLREOKFEYVYAOSLOVPYRNSLRKAHIGVFSFSTQAMMPYSVACFR 960
Db 840 EAIENFTVSVLREOKFEYVYAOSLOVPYRNSLRKAHIGVFSFSTQAMMPYSVACFR 899
Qy 961 FGATLVANPEPNPDVLLVSALVFGMAVGOVSTPAPDYAKAKVSAHVIMITEKSPLI 1020
Db 900 FGATLVANPEPNPDVLLVSALVFGMAVGOVSTPAPDYAKAKVSAHVIMITEKSPLI 959
Qy 1021 DSYSPHGLKPNTELEGNTEFNEVEFNTPTRPDIPIVLOGLSLEVKGGOTLALVSSCGCKST 1080
Db 960 DSYSPHGLKPNTELEGNTEFNEVEFNTPTRPDIPIVLOGLSLEVKGGOTLALVSSCGCKST 1019
Qy 1081 VVOLLERFYDPLAGSVLIDGKEIKHLNVOMLRAHLGIVSOEPIIDPCSAENIAAGDMSR 1140
Db 1020 VVOLLERFYDPMAGVTELDCKEIKHLNVOMLRAHMGIVSOEPIIDPCSAENIAAGDMSR 1079
Qy 1141 VVSHHEELMOAKKAPANIHHFETLPEKYNTRVGDKGTQLSGGOKORAIARALVROPHTLL 1200
Db 1080 VVSHHEELVRAKAKENIHOFIDSLPEKYNTRVGDKGTQLSGGOKORAIARALVROPHTLL 1139
Qy 1201 LDEATSAIDTESKVVQOALDKAR 1224
Db 1140 LDEATSAIDTESKVVQOALDKAR 1163
```

RESULT 6
Q60502 PRELIMINARY: PRT: 1169 AA.
AC Q60502;
DT 01-NOV-1996 (Tremblrel. 01, Created)
DT 01-NOV-1996 (Tremblrel. 01, Last sequence update)
DT 01-JUN-2002 (Tremblrel. 21, Last annotation update)

DE P-glycoprotein.
GN pgp-1.
OS Cricetus sp.
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Cricetinae;
OC Cricetus.
OC NCBI_TaxID=10031;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=LUNG;
RX MEDLINE=91154265; PubMed=1671863;
RA Devine S.E., Hussain A., Davide J.P., Melera P.W.;
RT "Full length and alternatively spliced pgp-1 transcripts in multidrug
KT resistant Chinese hamster lung cells."
RL J. Biol. Chem. 266:4545-4555(1991).
CC -1. SIMILARITY: BELONGS TO THE ABC TRANSPORTER FAMILY.
DR EMBL: M59254; AAA37005.1; -;
DR InterPro: IPR003593; AAA_Atpase.
DR InterPro: IPR001140; AbcTransporter.
DR InterPro: IPR003439; ABC_transporter.
DR Pfam: PF00664; ABC_membrane_2.
DR Pfam: PF00005; ABC_tran_2.
DR PRODOM: PD000006; ABC_transportr_2.
DR SMART: SM00382; AAA; 2.
DR PROSITE: PS00211; ABC_TRANSPORTER; 2.
DR ATP-binding; Transport.
KW SEQUENCE 1169 AA; 128939 MW; 72E25B7CE29DC185 CRC64;
SQ
Query Match 82.3%; Score 5333; DB 11; Length 1169;
Best Local Similarity 89.2%; Pred. No. 6.4e-301;
Matches 1043; Conservative 69; Mismatches 57; Indels 0; Gaps 0;
OY 113 MTTAAAYYSIGAGVLAAYIOVSEWCLAAGQILKIRKOPFHAIMROEIGFVHDVGE 172
DB 1 MTTAAAYYTGIGAGVLIVAVIQVSFWCLAAGQIHRIRKPFHAIMNOEIGFVHDVGE 60
OY 173 LNTRLTDVSKINTEGDKIGMEFHSIATPFTEIGVTRGKMLTLVLAISPVGLSAA 232
DB 61 LNTRLTDVSKINTEGDKIGMEFHSIATPFTEIGVTRGKMLTLVLAISPVGLSNG 120
OY 233 IMAKILSFTDKELLYAKAGAAVEVLAIRTVAIFGQKLELERYKNLEAKGIGIK 292
DB 121 IMAKILSFTDKELLYAKAGAAVEVLAIRTVAIFGQKLELERYKNLEAKRLGK 180
OY 293 KAITANISIGAFLLIYASALAFWGTSTVLSSEYSIGOVLTVEFSVLIGASIGQASP 352
DB 181 KAITANISIGAFLLIYASALAFWGTSTVLSSEYSIGOVLTVEFSVLIGASIGQASP 240
OY 353 SIEAFANRGAAYEIFKILDKPSIDSYSGSKHPPNIKONLEKKNVHFSYPSRKEVKIL 412
DB 241 SIEAFANRGAAYEIFKILDKPSIDSYSGSKHPPNIKONLEKKNVHFSYPSRKEVKIL 300
OY 413 KGLNKKVSGQTVLVNKGSGKSTVOLMQLYDPIDGAVCIDGODIRITNVRHIREIT 472
DB 301 KGLNKKVSGQTVLVNKGSGKSTVOLMQLYDPIDGAVCIDGODIRITNVRHIREIT 360
OY 473 GYVSEPELLEFATTAENIRYGRENVTMDEIEKAVKEANAYDFIMKLPNKEDLVGRGAQ 532
DB 361 GYVSEPELLEFATTAENIRYGRENVTMDEIEKAVKEANAYDFIMKLPNKEDLVGRGAQ 420
OY 533 LSGGKORITATARALVRNPKILLDEATSAIDTESEAVVVALDKARKGTTTIVIAHRIS 592
DB 421 LSGGKORITATARALVRNPKILLDEATSAIDTESEAVVVALDKARKGTTTIVIAHRIS 480
OY 593 TVRANADVLAGPDGYIVYKGNHDLMEKGIYFKLYMTQKGNIELENTGSKSBSDA 652
DB 481 TVRANADVLAGPDGYIVYKGNHDLMEKGIYFKLYMTQKGNIELENTGSKSBSDA 540
OY 653 LEMSPKSGSSLIKRRSRRSIIHAPQODRKLTGTEDLNENVPVSEFWRIKLKUNSTWY 712
DB 541 LEMSPKSGSSLIKRRSRRSIIHAPQODRKLTGTEDLNENVPVSEFWRIKLKUNSTWY 600
OY 713 FVVGIFCALINGIQAPAFSIIFSRILIGFTREDEDETRKONSNFSLVFLVGLISRTIF 772

DB 601 FVVGIFCALINGIQAPAFSIIFSRIVGVVFTRNTDDETRKRDNSLFLSLFLIGVISFTIF 660
OY 773 FLGGFTFEKAGEILTKRLRYVFRSMRLRQDSYFDDPKNTGTLTRTLANDAQVKAIG 832
DB 661 FLGGFTFEKAGEILTKRLRYVFRSMRLRQDSYFDDPKNTGTLTRTLANDAQVKAIG 720
OY 833 SRLAVITONIANLGTGIIISLIYGMOLTLTLLAIVPIAIVGVENKMKLSGALKDKREL 892
DB 721 ARLAVITONIANLGTGIIISLIYGMOLTLTLLAIVPIAIVGVENKMKLSGALKDKREL 780
OY 893 EGAGKIATEAENFRTVSLTRQKFEYMAQSLQVPRNSLKAHIFGVSESIPTAMMY 952
DB 781 EGAGKIATEAENFRTVSLTRQKFEYMAQSLQVPRNSLKAHIFGVSESIPTAMMY 840
OY 953 FSYACCFRGAIVLANEFMNRQDYLVSFAIVFGAMAVGVSSFPADYAKAKYASAHIT 1012
DB 841 FSYACCFRGAIVLANEFMNRQDYLVSFAIVFGAMAVGVSSFPADYAKAKYASAHIT 900
OY 1013 IIEKSPILDSYSPHGLAKNTELEGNTEVENYPTRPDIPLVLOGLSLEVKKQOTLALV 1072
DB 901 IIEKSPILDSYSTGLKNTLEGNKKEVEYENYPTRPDIPLVLOGLSLEVKKQOTLALV 960
OY 1073 SSGCGKSTVOLLERPYPLAGSVLIDGKEIKHNVOMLRAHIGIVSOEPILEDCSTAE 1132
DB 961 SSGCGKSTVOLLERPYPLAGSVLIDGKEIKHNVOMLRAHIGIVSOEPILEDCSTAE 1020
OY 1133 IAYGDSRYVSHETMOAKKEANIHFEPTLEPEKYNFVGRGKQVLSGGOKORITARAL 1192
DB 1021 IAYGDSRYVSHETMOAKKEANIHFEPTLEPEKYNFVGRGKQVLSGGOKORITARAL 1080
OY 1193 VROPHTLLDEATSAIDTESEKRYVOALDKAREGTCIVIAHRISTONADLIYVFPNG 1252
DB 1081 VROPHTLLDEATSAIDTESEKRYVOALDKAREGTCIVIAHRISTONADLIYVFPNG 1140
OY 1253 VKEHGTHOOLAKQGYFSMVSVOGAKR 1281
DB 1141 VKEHGTHOOLAKQGYFSMVSVOGAKR 1169

RESULT 7
ID 08RA27 PRELIMINARY; PRT; 1275 AA.
AC 08RA27;
DT 01-JUN-2002 (TREMblrel. 21, Created)
DT 01-JUN-2002 (TREMblrel. 21, Last sequence update)
DT 01-JUN-2002 (TREMblrel. 21, Last annotation update)
DE ATP-binding cassette protein B1b.
OS Rattus norvegicus (Rat).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.
OX NCBI_TaxID=10116;
RN [1]
RP SEQUENCE FROM N.A.
RC SPRALIN-SPRAGUE-DAWLEY; TISSUE=LIVER;
RA Yabuuchi H., Ishikawa T.;
RL Submitted (MAR-2002) to the EMBL/GenBank/DBJ databases.
DR EMBL: AY082609; AAJ92458.1; -;
KW ATP-binding.
SQ SEQUENCE 1275 AA; 141222 MW; 70D0E5F6A0E0E19D CRC64;
Query Match 81.6%; Score 5286.5; DB 11; Length 1275;
Best Local Similarity 80.3%; Pred. No. 3.7e-298;
Matches 1029; Conservative 130; Mismatches 114; Indels 9; Gaps 5;
OY 1 MDPEGRGSAEKNFWMKSKSKNEKKEKPPYSTRAMPYRNWLDRLMYLVGTMAAII 60
DB 1 MEPEGLNGTADKKNFSKSKSKK--EKERKPAVGIFGMFRIVADWLDKLCMALGTAAII 58
OY 61 HGAALPLMLVAFGMNTOSEFANAGISRNKTFPVIIINESITNTQHTI-NHLEEDMTYAY 119
DB 59 HGTILPLMLVFGYMTDSFTA---ETRIPLSVYNGEINSTQTVSDSLEEDMAYAY 115

```

QY 120 YSGAGVLAAYIYQVSWFMCILAAGROILKTRKOFPHAIINBOIGMDVHDVGLNRLTD 179
D 116 YTGAGVLAAYIYQVSWFMCILAAGROILKTRKOFPHAIINBOIGMDVHDVGLNRLTD 175
QY 180 DVKINGINGIDKIGMFHSHATPTFTGFTVGTGKMLTVIIAISPVLGSAIAWAKITS 239
D 176 DVKINGINGIDKIGMFHSHATPTFTGFTVGTGKMLTVIIAISPVLGSAIAWAKITS 235
QY 240 SFTDKELATAKAGAAVEEVLAAIRTVIAPGCKKLEKRYNNKLEAKGIGIKKAITANI 299
D 236 SFTDKELATAKAGAAVEEVLAAIRTVIAPGCKKLEKRYNNKLEAKGIGIKKAITANI 295
QY 300 SIGAFTLLIYASYALAEMWGTSLVLSSEYSIGCVLTFEFSVLIGAFSIGQASPSIEARAN 359
D 296 SIGAFTLLIYASYALAEMWGTSLVLSSEYSIGCVLTFEFSVLIGAFSIGQASPSIEARAN 355
QY 360 ARGAAVEIFKIDNKPSISYSKSGKPDNIKGNLEFKNVHESYPSRREKVKILKGLNKKV 419
D 356 ARGAAVEIFKIDNKPSISYSKSGKPDNIKGNLEFKNVHESYPSRREKVKILKGLNKKV 415
QY 420 QSGQVAVLVNSGCGKSTTVQLMORLYDPTDGMVCIIDODRTINVRRLREITGVSOEP 479
D 416 KSGQVAVLVNSGCGKSTTVQLMORLYDPTDGMVCIIDODRTINVRRLREITGVSOEP 475
QY 480 VLEFATIAINIRYGRBNVTMDIEKAVKAAVADFIKMLPKNFPTLVGERGAQSGGOKO 539
D 476 VLEFATIAINIRYGRBNVTMDIEKAVKAAVADFIKMLPKNFPTLVGERGAQSGGOKO 535
QY 540 RIAIAALVRNPKILLDEATASLTDESEAVOVALDKARGRITVIAHRLSTVRNADY 599
D 536 RIAIAALVRNPKILLDEATASLTDESEAVOVALDKARGRITVIAHRLSTVRNADY 595
QY 600 IAGFDGVYVEKGNDELAKKEGIYFKLVMTQRTGNEIETENATGESKSDALEPKD 659
D 596 IAGFDGVYVEKGNDELAKKEGIYFKLVMTQRTGNEIETENATGESKSDALEPKD 655
QY 660 SSSSLIKRSTRIRSHAPOGODRKIGKEDENAVPVYSWRILKLNSTWPFYVVGIFC 719
D 656 SSSSLIKRSTRIRSHAPOGODRKIGKEDENAVPVYSWRILKLNSTWPFYVVGIFC 714
QY 720 AIIINGLOPAFSIFSRILIGITRDEDEPETRKONNSFVLIVLIGITSIFTEGTF 779
D 714 AIIINGLOPAFSIFSRILIGITRDEDEPETRKONNSFVLIVLIGITSIFTEGTF 774
QY 780 GAGELILKRLRYWVFSMLRQDYSPDDPKNTGALTIRLANDAAQVKAIGSRLAVT 839
D 774 GAGELILKRLRYWVFSMLRQDYSPDDPKNTGALTIRLANDAAQVKAIGSRLAVT 834
QY 840 QNANLGTGIIISLIYQMOLILLLAIIVPIIATAGVEMKMSGQALKKKELEGAKTA 899
D 834 QNANLGTGIIISLIYQMOLILLLAIIVPIIATAGVEMKMSGQALKKKELEGAKTA 894
QY 900 TEALENRTVYSLTRECKFEYMTAQSLOVPYRNSLRKAHIFGVSESIQAMMYFSYAGCF 959
D 894 TEALENRTVYSLTRECKFEYMTAQSLOVPYRNSLRKAHIFGVSESIQAMMYFSYAGCF 954
QY 960 RFAGVILVARNLTFENMVLVFSNAVFGAAGAGTSSFADYAKAKVASHIIRIETKIDE 1019
D 954 RFAGVILVARNLTFENMVLVFSNAVFGAAGAGTSSFADYAKAKVASHIIRIETKIDE 1014
QY 1020 IDSISPGKLPNTLEGNTFNEVFNPTRPDIPLYQGLSLEVKKGOTLALVGSSGCKS 1079
D 1014 IDSISPGKLPNTLEGNTFNEVFNPTRPDIPLYQGLSLEVKKGOTLALVGSSGCKS 1074
QY 1080 TVVOLLERFDPPLAGSVLDGKEIKILANOWMLRAHLGIYSQEPILFDCSIAENIAYGNS 1139
D 1074 TVVOLLERFDPPLAGSVLDGKEIKILANOWMLRAHLGIYSQEPILFDCSIAENIAYGNS 1134
QY 1140 RVVSHHEIQAQKAAKANIHFIELEPEKYNTRYGDKGTOLSGOKORIALAALVRQPHIL 1199
D 1134 RVVSHHEIQAQKAAKANIHFIELEPEKYNTRYGDKGTOLSGOKORIALAALVRQPHIL 1194
QY 1200 LDEATASLTDESEKVVQALDKARGRITVIAHRLSTIQNADLIVFGNGKVEHGT 1259

```

```

D 1195 LDEATASLTDESEKVVQALDKARGRITVIAHRLSTIQNADLIVFGNGKVEHGT 1254
QY 1260 QQLAOKGIFYSMVQAGAKR 1281
D 1255 QQLAOKGIFYSMVQAGAKR 1274

RESULT 8
093437 PRELIMINARY: PRT: 1288 AA.
AC 093437;
DT 01-NOV-1998 (TREMblrel. 08, Created)
DT 01-DEC-2001 (TREMblrel. 19, last sequence update)
DT 01-JUN-2002 (TREMblrel. 21, last annotation update)
DE ABC transporter protein.
GN CMDR1.
OS Gallus gallus (Chicken).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Archosauria; Aves; Neognathae; Galliformes; Phasianidae; Phasianinae;
OC Gallus.
OX NCBI_TaxID=9031;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=INTESTINE;
RX MEDLINE=9209805; PubMed=10195430;
RA Edelman H.M.L., Duchek P., Rosenthal F.E., Foeger N., Glackin C.,
RA Kane S.E., Kuchler K.,
RT "Cmdr1, a chicken P-glycoprotein, confers multidrug resistance and
RT interacts with Estradiol."
RL Biol. Chem. 380:231-241(1999).
CC -! SIMILARITY: BELONGS TO THE ABC TRANSPORTER FAMILY.
DR EMBL: AJ009799; CAA0835.1; -.
DR HSSP: P13569; INBD.
DR InterPro: IPR003593; AAA_ATPase.
DR InterPro: IPR001140; ABCtransportTM.
DR InterPro: IPR003439; ABC_transport.
DR InterPro: IPR000504; RNA_rec_mot.
DR Pfam: PF00664; ABC_membrane; 4.
DR Pfam: PF00005; ABC_tran; 4.
DR ProDom: PD00006; ABC_transport; 2.
DR SMART: SM00382; AAA; 4.
DR PROSITE: PS00211; ABC_TRANSPORTER; 4.
DR PROSITE: PS00030; RRM_RNP_1; UNKNOWN_2.
KW ATP-binding; Transport.
FT CHAIN 2 1288
SQ SEQUENCE 1288 AA; 141917 MW; CB258A5F2826DB8C CRC64;

Query Match 72.0%; Score 4660.5; DB 13; Length 1288;
Best Local Similarity 71.2%; Pred. No. 8.1e-262;
Matches 919; Conservative 160; Mismatches 167; Indels 45; Gaps 12;

QY 2 DPBGGRKGALEKKNFKWKKSKKNEKKEKKP-TVSTFAMFRYSNMLDRLYMLVGTMAAI 60
D 27 DPDEKKG-----KKKKKKPQVPSPLAFRRSSCDKLMIFGSLAIA 71
QY 61 HGAALPLMLVFGNMTDSFANAGISRNKTPPYIYNSTNNTOHNLHEEMTYAVYY 120
D 72 HGSISPLAMTIFSDMTDSFTSGMTN-----ITGSSGNSADVFNKLEEMTRYAVYY 126
QY 121 SGIAGVLAAYIYQVSWFMCILAAGROILKTRKOFPHAIINBOIGMDVHDVGLNRLTD 180
D 127 SAIAAVALVAAYIYQVSWFMCILAAGROILKTRKOFPHAIINBOIGMDVHDVGLNRLTD 186
QY 181 VSKINGINGIDKIGMFHSHATPTFTGFTVGTGKMLTVIIAISPVLGSAIAWAKITS 240
D 187 VSKINGINGIDKIGMFHSHATPTFTGFTVGTGKMLTVIIAISPVLGSAIAWAKITS 246
QY 241 FTDKELATAKAGAAVEEVLAAIRTVIAPGCKKLEKRYNNKLEAKGIGIKKAITANIS 300
D 247 FTDKELATAKAGAAVEEVLAAIRTVIAPGCKKLEKRYNNKLEAKGIGIKKAITANIS 306
QY 301 IGAFTLLIYASYALAEMWGTSLVLSSEYSIGCVLTFEFSVLIGAFSIGQASPSIEARANA 360

```

```

Db      :|||||
307 MGAFFLIYASALAEWYGTLLHLNLEISGNLVTFEFSVLGAGSISGOTAFASIAFANA 366
Qy      361 RGAAYEIFKIDNKPISIDSYSGKHPRDNKGNLEFRNHFSTPSRKEVKIILKGLNKVY 420
      |||||
Db      367 RGAAYAFINIIDNEPIDSYSDAGHKPDHDKIKLEFONFENFNSPRDVEILKGLNKVY 426
Qy      421 SGOTVALVNSGCGKSTTVOLMORLYDPDGMWCIDGODIRITNVNHLREITGVNSOEFP 480
      |||||
Db      427 CGOTVALVNSGCGKSTTVOLIQREFIDPKEGTTTIDGOLSLNVYKLEIIGVNSOEFP 486
Qy      481 LEATTIAENIRYGRENTYMDIEKAVKANAYDFIMKLPNKFDVLVGERGQALSGGQOKR 540
      |||||
Db      487 LEATTIAENIRYGRREDYMEIERATEKANAYDFIMKLPKFEFTVYGERGQALSGGQOKR 546
Qy      541 IAIARALVRNPKIILLDEATSALDTESEAVVOVALDKARKGRTTYIAHRLSTYVRNADY 600
      |||||
Db      547 IAIARALVRNPKIILLDEATSALDTESEAVVOVALDKARKGRTTYIAHRLSTYVRNADY 606
Qy      601 AGPDGVIVKGNDELKKEGIYFKLVMTQTRNGNEIELENATGESKESDALEMSPKDS 660
      |||||
Db      607 AVFESGVITTEQGNHSLIEKKGIYKLVNMQT-----IETEDPSE--KSEN---AVSYKRS 658
Qy      661 GS-----SL---IKRSTRSIIH---APGODRKLGTKEPDLNVPVPSFWMIILKNSF 708
      |||||
Db      659 GSQSNLDESLKELRRGSTRSMKKPKPEPNDTDEK--GSSPD--BELPPVSFLKMLKLNK 715
Qy      709 EMPYVVGIFCAIINGLOPAFSITFSRIIGITREDEPPTKRONSMFSLVLGIIIS 768
      |||||
Db      716 EMPYVVGIFCAIINGLOPAFSITFSRIIGITREDEPPTKRONSMFSLVLGIIIS 774
Qy      769 FTFEFLQGTFFKAGEIILTKRLMYVFRSMRLQDVSWFDDPKNTTGALTTPFLANDAQVK 828
      |||||
Db      775 FTFEFLQGTFFKAGEIILTKRLMYVFRSMRLQDVSWFDDPKNTTGALTTPFLANDAQVK 834
Qy      829 GAIGSLAVITTONIANLGGIIISLIYGMOLTLLALLVPIITAIAGVYEMKMLSGQALKD 888
      |||||
Db      835 GAGVGRALIAQNIANLGGIIISLIYGMOLTLLALLVPIITAIAGVYEMKMLSGQALKD 894
Qy      889 KKELEGAGRIATEALENFTVSVLTREOKFEYVYASIQVPRNSLKRKAHIFGVSFSTIQ 948
      |||||
Db      895 KIELEAGRIATEALENFTVSVLTREOKFEYVYASIQVPRNSLKRKAHIFGVSFSTIQ 954
Qy      949 AMMYFSTACFRGATLVANNEFNPQDVLVSAIVFGMAVGOVSSFPADYAKAKYSAA 1008
      |||||
Db      955 AMMYFSTACFRGATLVANNEFNPQDVLVSAIVFGMAVGOVSSFPADYAKAKYSAA 1014
Qy      1009 HVTIIEKSPIDISYSPHGLKPTLEGNVTFNEVVFENYFTRPDIPYLOGSLEVKKQTL 1068
      |||||
Db      1015 HVEVLFRRVPRIDSYREDGKPEKFGNTRIKDVKNFNYPNRBEVKILOGLANLVEKGETL 1074
Qy      1069 ALVSSGCGKSTVOLLERYDPLASVYLDGKEIKHLNWOVLAHLIGVSOEPIIFDCS 1128
      |||||
Db      1075 ALVSSGCGKSTVOLLERYDPLASVYLDGKEIKHLNWOVLAHLIGVSOEPIIFDCS 1134
Qy      1129 IAEVIAVGDNSRVVSHHEIQAKAEANIHFETLEPEKYNTRYVGDGQOLSGGQOKRIAI 1188
      |||||
Db      1135 IAEVIAVGDNSRVVSHHEIQAKAEANIHFETLEPEKYNTRYVGDGQOLSGGQOKRIAI 1194
Qy      1189 ARALVRPHILLDEATSALDTESEAVVOVALDKARKGRTTYIAHRLSTYVRNADY 1248
      |||||
Db      1195 ARALVRPHILLDEATSALDTESEAVVOVALDKARKGRTTYIAHRLSTYVRNADY 1254
Qy      1249 ONGKVEHGTTHOOLLAQGIYFNSVVOAGA 1279
      |||||
Db      1255 ONGKVEHGTTHOOLLAQGIYFNSVVOAGA 1285

```

RESULT 9
 Q91586 PRELIMINARY; PRT; 1287 AA.
 ID Q91586
 AC Q91586
 DT 01-NOV-1996 (Tremblrel. 01, Created)

DT 01-NOV-1996 (Tremblrel. 01, last sequence update)
 DT 01-JUN-2002 (Tremblrel. 21, last annotation update)
 DE Multidrug resistance protein.
 GN XEMDR.
 OS Xenopus laevis (African clawed frog).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Amphibia; Batrachia; Anura; Mesobatrachia; Pipidea; Pipidae;
 OC Xenopodidae; Xenopus.
 NC NCBI_TaxID=8355;
 RN [1]
 RP SEQUENCE FROM N.A.
 RX MEDLINE=95322451; PubMed=7599185;
 RA Castillo G., Shen H.J., Horwitz S.B.;
 RT "A homologue of the mammalian multidrug resistance gene (mdr) is
 functionally expressed in the intestine of *Xenopus laevis*."
 RL Biochem. Biophys. Acta 1262:113-123(1995).
 CC -1- SIMILARITY: BELONGS TO THE ABC TRANSPORTER FAMILY.
 DR EMBL; U17608; AAA75000.1; -;
 DR InterPro; IPR003593; AAA_ATPase.
 DR InterPro; IPR001140; ABCtransport.
 DR InterPro; IPR003439; ABC_transport.
 DR Pfam; PF00064; ABC_membrane_2.
 DR Pfam; PF00005; ABC_tran; 2.
 DR ProDom; PD00006; ABC_transport; 2.
 DR SMART; SM00382; AAA; 2.
 DR ProSITE; PS00211; ABC_TRANSPORTER; 2.
 DR ATP-binding; Transport.
 KW SEQUENCE 1287 AA; 141505 MW; 06E95073C5711415 CRC64;

Query Match 68.4%; Score 4430; DB 13; Length 1287;
 Best local similarity 67.6%; Pred. No. 2e-248;
 Matches 871; Conservative 172; Mismatches 211; Indels 34; Gaps 11;

```

Qy      2 DEGGKSGAENKFNKMKKSKKKEKPTVSTFAMFRYSNKLDFLYLVGTMAIIT 61
      |||
Db      21 DPNNSK--EKGF--SKFKKKKTEKPKYGVFTFWFRYSSTSDKMLFGTISLAH 76
Qy      62 GAALPLMLVFGNMTPDSRANAG--ISRKTEPVIIINESTINNTQHFVNHLEEMTAYY 119
      |||||
Db      77 GAALPLMLVFGEMTDFVNGVDVTFNFWESKINAS-----RELQGGTAYY 127
Qy      120 YSGIGGVVAAYIQVSEWCLAGROIILKIRKOPHAIMROEIGMPDVHVGELNTLTD 179
      |||||
Db      128 YSGIGGVVAAYIQVSEWCLAGROIILKIRKOPHAIMROEIGMPDVHVGELNTLTD 187
Qy      180 DVSKINEIGIDKIGFFHSIAFTFGFVGFGRKGLVILATSPVLSAIAIMAKILS 239
      |||||
Db      188 DVSKINEIGIDKIAMLQSLTTLVTFIIGFTKGMKTLVWGALSPINGLSAIAIMAVLS 247
Qy      240 SEFDKELAVAKAGAAVEVLAIRTVIAFGGQKKELERYNKNLEAKGIGIKKAITANI 299
      |||||
Db      248 AFTNKEIKAKAGAAVEVLSIRTVFAFGQNKIEHRYEKNLEDAKKIGIKKAITANI 307
Qy      300 SIGAPELLIYASALAEWYGTLSVLSSEYSGOVITVFPFSLVIGAFSGOASPIEAFAN 359
      |||||
Db      308 SIGAPELLIYASALAEWYGTLSVLSSEYSGOVITVFPFSLVIGAFSGOASPIEAFAN 367
Qy      360 ARGAAVTEFTIINKKSIDSYSGKHPRDNKGNLEFRNHFSTPSRKEVKIILKGLNKV 419
      |||||
Db      368 ARGAAVTEFTIINKKSIDSYSGKHPRDNKGNLEFRNHFSTPSRKEVKIILKGLNKV 427
Qy      420 QSGOTVALVNSGCGKSTTVOLMORLYDPDGMWCIDGODIRITNVNHLREITGVNSOEFP 479
      |||||
Db      428 QSGOTVALVNSGCGKSTTVOLIQREFIDPKEGTTTIDGOLSLNVYKLEIIGVNSOEFP 487
Qy      480 VLEFTTIAENIRYGRENTYMDIEKAVKANAYDFIMKLPNKFDVLVGERGQALSGGQOKR 539
      |||||
Db      488 VLEFTTIAENIRYGRENTYMDIEKAVKANAYDFIMKLPNKFDVLVGERGQALSGGQOKR 547
Qy      540 IAIARALVRNPKIILLDEATSALDTESEAVVOVALDKARKGRTTYIAHRLSTYVRNADY 599
      |||||
Db      548 IAIARALVRNPKIILLDEATSALDTESEAVVOVALDKARKGRTTYIAHRLSTYVRNADY 607

```

```

QY 600 IAGDGVIVEKGNHDELMKNGIYFKLVMTOT-----RGNETELENAEGSESDALEM 655
Db 608 IAGDNGVIVEQSGHKEMLMGVYVNLVLTQVETSKTEDELTTHYEKK-----I 660
QY 656 SPDSQSSILKRSTRSHHA--PGQDRKLTKE-DLNEENPVPSFMRILKNKPEMPY 712
Db 661 PVHTHSNLTNRKSSRNTKSKYPTEDKEVDEEKKKEEGSPVPSFKVKNLKNKPEMPY 720
QY 713 FVVGICAIINGLQAPAFISIRIIGFTPREDEPRK-RONSNNESVLFVIGISFIT 771
Db 721 FVVGICAMINGATOPAFIISRIIGVFA---GVVSQMRSESSVSLFALGAGSPIT 777
QY 772 FFLQGFTEPKAGEIILTKRLRYVFSMLRODVSWFDDPKNTGALTPELADDAOVGAI 831
Db 778 FFLQGFTEPKAGEIILTKRLRYVFSMLRODVSWFDDPKNTGALTPELADDAOVGAI 837
QY 832 GSRLAVITONIANLTGIIISLIYQMLTLLIIVPIIAAGVYEMKMSGALKDKKE 891
Db 838 GTRMLAQNVANLTGIIISLIYQMLTLLIIVPIIAAGVYEMKMSGALKDKKE 897
QY 892 LEGAKIATLEAENFRVYVSLTREQKEVYMAQSLQVYRNSLRKAHIFGVSPSIQAM 951
Db 898 LEKAKISTDAVLRVYVSLTREQKEVYMAQSLQVYRNSLRKAHIFGVSPSIQAM 957
QY 952 YFSVAGCFR-FGAYLVANEFNFDVLLVFSATYFGAMANGVYSSAPDYAKAKVSAHY 1010
Db 958 VILCQWFSVAGLVLVEGLMKLDEVEFVSSATYLVGLMAGVOTSSAPDYAKAKVSAHY 1017
QY 1011 IMIEKSLDISYSPHGLKPNTELEGNVTFNEVYVPTRPDIPLYGLSLVYKQGTAL 1070
Db 1018 FSLLEFRPOIDYSVDOGEKPKNSGVNVFKGVNFNPTRPDIPLYGLSLVYKQGTAL 1077
QY 1071 VGSSCGKSTVYVOLLERFVYPLAGSVLIDGKEIKHLNVOMLRHAGIYQOEPILEDCSIA 1130
Db 1078 VGSSCGKSTVYVOLLERFVYPLAGSVLIDGKEIKHLNVOMLRHAGIYQOEPILEDCSIA 1137
QY 1131 ENIAYGNSVSVSHEETMAKAEANIHFELETPERYNRYVDKDTQSLSGGKOKIATAR 1190
Db 1138 DNIAYGNNRKYQOELETAKEANIHFELETPERYNRYVDKDTQSLSGGKOKIATAR 1197
QY 1191 ALVROPHILDEATSAIDTESKRYQVQALDKARSGRCIYAHRLSTIONADILVYFON 1250
Db 1198 ALIRKRIILDEATSAIDTESKRYQVQALDKARSGRCIYAHRLSTIONADILVYFON 1257
QY 1251 GKVEHGTHQOLLAOKGIYFSMVYQAG 1278
Db 1258 GKVEHGTHQOLLAOKGIYFSMVYQAG 1285

```

RESULT 10
090235 PRELIMINARY; PRT; 1348 AA.

```

AC 090235;
DT 01-DEC-2001 (TREMBlrel. 19, Created)
DT 01-DEC-2001 (TREMBlrel. 19, Last sequence update)
DT 01-JUN-2002 (TREMBlrel. 21, Last annotation update)
DE Bile salt export pump.
OS Raja erinacea (little skate).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Chondrichthyes;
OC Elasmobranchii; Squala; Hypnosquala; Pristioraja; Batoidae;
OC Rajiformes; Rajidae; Raja.
NCBI_taxid=7782;
RX
RP SEQUENCE FROM N.A.
RA MEDLINE-21340059; PubMed-11447010;
RA Cai S.Y., Wang L., Ballatori N., Boyer J.L.;
RT "Bile salt export pump is highly conserved during vertebrate evolution
RT and its expression is inhibited by pfc type II mutations.";
RL Am. J. Physiol. Gastrointest. Liver Physiol. 281:G316-G322(2001).
DR EMBL; AF367243; AAK52958.1;
DR InterPro; IPR001140; ABCtransp17M.
DR InterPro; IPR003439; ABC_transp17.
DR Pfam; PF00664; ABC_membrane; 2.

```

```

DR Pfam: PF00005; ABC_tran; 2.
DR ProDom: PD000006; ABC_transp17; 2.
DR PROSITE; PS00211; ABC_TRANSPORTER; UNKNOWN.1.
SQ SEQUENCE 1348 AA; 148648 MW; 81906F641654CDD CRC64;

```

```

Query Match 51.4%; Score 3331; DB 13; Length 1348;
Best local similarity 49.8%; Pred. No. 1.3e-184;
Matches 661; Conservative 253; Mismatches 355; Indels 58; Gaps 10;

```

```

QY 1 MDEGGRK-GSAEKFWKMKSKKN-EKKEKPPVSTFAEFRYSNMILRLYLVGMMA 58
Db 27 IDGPGGRKYVYNGVYVPSKXDNENLSKDDRIKIGFQFRASCVEYTLVWGVCA 86
QY 59 ITHGALPLMMVPEGMNDSPFANAGISNKKTFPIYINESTNNTOHFIN----- 107
Db 87 LHHGAQRAVLVFLDLTFDFAVDIELQELKDT--RKHCNNNTIHWNGSEYLNQNER 144
QY 108 -----HLEEMTYAVYVSGIGAGVLYVAAIYQVSPCLAAQRQILKIRQFHAIKROE 161
Db 145 MSCGILNTEKEKMTMAYVYVIGGCVLLIGYFOICFWYTAARQOTIRAYRQIMRE 204
QY 162 IGMFVHVHVGELNRLDDVSKINIGDKIMGFHSHIAFTFGFTVGTGRKTLVLI 221
Db 205 MGVFDCNSVGLNTRMSDDINKINDAIDAVGVGIFQRFSTFVSGFLMGFPVNGKLTLLI 264
QY 222 AISPVLGSAIIMAKIISFTDKELLAYAKAGVAEVLAAIFTYVAFGQKKELERYK 281
Db 265 AVSPLIGIGALMALVAVRLIGIKRAYAKAGVADEVLSIRTYVAFSGEKREYDRDR 324
QY 282 NLEAKIGIKKATIANISIGAFLLIYASALAFYGTSLV--SSYSIGVLIYFVS 340
Db 325 NLVFAQWQGRKRMKMGVFTGYWMIIFCCYALAFYGSKLVEQNEYRPGLLQVFLCV 384
QY 341 LIGAFSIGQASPSIEFAANARCAAYELFRTIDNKPISIDYSKSGKPNKGNLEFKNV 400
Db 385 LVAAAMNGQASPOLAEFAFGSGRAVAAYFETIDNEPDCMSGVTLNKKVGDIEFHNVT 444
QY 401 FSVPSREKRIILKGLNLYQSGQVVALVGNSSCGKSTVQLMQRLYDPDGVNCIDGDI 460
Db 445 FNPSPRVDKTLDRISMVAKGETTAFVGPSSGSGSTAVELLQRYDYDKQGMVILGDHI 504
QY 461 RTINVHLREITGVYQOEVLFTATTIEMIRYGRNVYMDLEIKAVKANAAYDFIMKLNV 520
Db 505 RSLNIQWLSLIGIYQOEVLSTIIEENIRYRGKLNQIVYAAKANAAYDFIMDLE 564
QY 521 KEDTVYGERGQOLSGQORAIARALVNRKILILDEATSAIDTESAPVQVVALDKAR 580
Db 565 KENTLVGEGGMSGGQORAIARALVNRKILILDMATSAIDNESALVQALDKVR 624
QY 581 GRTTIVAHRLSTVNAVVIAGFDGVIVEKGNHDELMKNGIYFKLVMTOTRGNTELE 640
Db 625 GRTTISIAHRLSTVNAVVIITGFEGRAVERGHAELLERGIFTLVLTOTGEOALHE 684
QY 641 NATGSKESDALENKSPDSQSSILKRSTRSHHAPOGDRK----- 683
Db 685 KA---RQVNGAIEDGASEK-RQLIRGSSNAYVSRIRHRSSQVSEVLSLSCGDVA 739
QY 684 -----LGTK-----EDLNEENPVPSFMRILKNSTEMPYVVGIFCAIINGLQAPAF 730
Db 740 SAVRTPSISLDEDEKDYEEBESLEPAPVSRILKYNSEPVMYLFSLCAAVGVGNPIY 799
QY 731 SIISRIIGFTPREDEPRKONSNNESVLFVIGISFTPELOGTEGKAGEIILTKRL 790
Db 800 ALFSQILGTFSL-QNEEIKINQINICLFVYVGLVSLTFOLOSIFPAKSGILLTREL 858
QY 791 RYVFRSMLRODVSWFDDPKNTGALTPELADDAOVKAISRLAVITONIANLTGII 850
Db 859 RKLQFQMLMQEIGMFDDBKRNNSPGLTTRPADASQVQATQIGMIVNSITNIGVSLI 918
QY 851 ISLIYQMLTLLIIVPIIAAGVYEMKMSGALKDKKELEGKATATAIEKFRVYV 910
Db 919 IAFYFSKTLTIVLCPFLPALGALQAMRLTGFANQDKALEAAGUISSEALSIRITIA 978

```


SEQUENCE 1294 AA: 142490 MW: DB8ED2602FAA84DF CRC64:
Query Match 44.1%; Score 2857; DB 5; Length 1294;
Best Local Similarity 45.6%; Pred. No. 4.1e-157;
Matches 579; Conservative 239; Mismatches 422; Indels 30; Gaps 9;
22 SKNKKKKKPTVSTFAMFRYSNMIDRLYMGTAIIHGALPLMLVFGWMTSFPAN 81
17 SEKKKAPPPKISTIFOLRYSTVDRLMAVGIISGATGVLPLMSIIMGVNSQFV 76
82 AGISRRKTPPVIIINESTNNTQ-----HFINHEEMTYAYYYSGAGVLAAYI 136
77 LG-----TIFDPNSTASEKAAAEFSHEVIONCKIYY-----LGGIFPAGFLQAS 125
137 FMCIAAGROILKIRKOFHAIHROEIGMVDVHDELNTRLTLDVSKINIGTKIMFF 196
126 CFNVICKLSNRRRPFHVSVMNOELIAYDKNTSGTLNKLFDNLERREGTGKVLAF 185
197 HSIATFPTGFTVGTGKMKLITVILAIISPVIGLSAIIWAKILISPTDELLAVKAGAVA 256
186 QMAAQFTGFAVAFATYDMLTLTLMSSLSPPMICGLFLAKLALATATAEAKQYAVAGGIA 245
257 EEVLAIRITVIAFGCGKLEERYKNLEAKIGIKAITANISGAAPFLIYASYALAF 316
246 EEWLTSTRTVIAFNGEYECRKREDALHGGKTKIGIKSFLIAGLAGLAFVLIYASYCLAF 305
317 WYGTSLVLSSEYSIGOVLYAFPSVLIGAFSISGQSPSIEAFNARGAAYEIRKIDNRP 376
306 WVGTFNVYSGRLSGTIVLYFVSVMGSMALGQAGQGFATIGALGAASLKEVIDRIPE 365
377 IDSYSKSHKPNKIKGNLEFNHFSYSRKEVYKILKMLNLYQSGYVALYVNSGCCGS 436
366 IDATSTEGOTPEKISGRISVKNVKEFTYPTRADYKILKGVSLDAPQGYALVSGSCGS 425
437 TTVOIMQRLYDPTGMCVITDODIRITNVRHLREITGVSEOEVLPAATTAENIRYGEN 496
426 TIITQLQRYNPDAGCIIIDIPEDENIKYLRQLVGVSEOEPLNFTSIEQIRYGRSD 485
497 YVMDELKRYKRNANVDFIMLPKPFDTLVGERGAQISGGOKORIALARALVNPILL 556
486 VSEDIARLAKRNADFKTFPEBLNLTVDGVOVSGGOKORIALARALVNPILL 545
557 DEATSAIDTSEEAVALDRAKRCRTTIVIAHRLSTVRNADYVAGDDVYIEGNHDE 616
546 DEATSAIDTSEESIVSALENASRGRTTIVIAHRLSTVRNADKIYKAKQWMEVGHET 605
617 LMEKGYFELVYMOFRGNEIELENATGESKESDLEMSPKDSGLIKRRSTR----- 671
606 LIEOKGLIHELVAQYFA---DVDDKPKKKEAERMSRQTSQKRGSVNFKTOESQVDEKP 662
672 RSHAPQODRKLQ--TKEDLNENPVPSFMKILKLNSTEMPYFVVGIFCAIINGLOPA 729
663 GAPAPAPAEKEIKRLKLEEGAVKANLFILKRAPEMITYFFALIIALLIGAVMFA 722
730 FSIIFSLITIGTREDDEPETKRONSNMFSVLVLVGIISFTIFLQGTFSKAGEIILTK 789
723 FSLFSSQIINVS- NPDRODKKDGHPALMFLVLAAYGTSMLQCSLFGAARLIMR 781
790 LRYWFRSMLRQDVSWFDPKNTTGALTITRLANDAQVAGVGLAVTQVIANIIGTGI 849
782 IRSKYRRLVLRQDATYFDMKHSFGRTITRLATDANINISALDIYRIGSFNMAIABVGG 841
850 IISLIGYQOLLLLATIAPVPIIAGVEMKMLSGQALKDKKELEGAKIATPAIENRTV 909
842 GIAFYGGOMAPLVMAIIFPEPAVGOALMKMYHGSATSDAKEMENKATAMAEINIRIV 901
910 VSLTRQKFEYNTAQSLOVYRNSLRKAHIFGVSTIQAMMFSAVSGEFGCAIYV--A 967
902 QALTITQTKLYNFCSHLDAPHGINSIKAIIRGLIYGFANSIDQFTYAAAFRGCLPIIFPK 961
968 NEPNANFOVLLVFSALVIGAMAVGOVSEFADYKAKAVSAAHYIMTIESSPLIDYSYPHG 1027
962 NYLMEPENLVRLVLAISFSFGTIGFASYPPEYIKATFAAGLIFNMLEEPRIDGMTSSG 1021

1028 LKPNTEGNTVENEVNPTRPDIPVQGLSLEVKKGQTLAVYSSCGCKSTVOLLER 1087
1022 TYQO-LSGEYKLNKVFERRPERPAVPILOGLVNHKKPQOTLALVPSGCKSTYISLLER 1080
1088 FYDPLAGSVLIDCKETKHLNVOMLRAHGLIVSQEPILFDCSIAENIAYGDNRSVSHET 1147
1081 LYPLEGAIVVNDNLRQONPKHLKHLIALVSOEPIILFDTSIRENIYVGLQPGETHEQI 1140
1148 MQAAKEANIHFETLEPEKYNTRVGDKTQLSGQOKORIALARALVROPHILLDEATSA 1207
1141 ETACSKANIKHFIDELPDGEYRVEGEGTQLSGQOKORIALARALIRPKILLDEATSA 1200
1208 LDTSEKVOQALDKAREGRTCIIVAHRLSTIONDLIVFPQNGKVEGTHQOOLLAOK 1267
1201 LDTSEKVOQALDAAKDRCTIVAHRLSTIYNACCIWVKNQGVBCGTHNELIARG 1260
1268 IYFSKAVVOA 1277
1261 AVFALTKROS 1270
RESULT 14
Q8T9W5 PRELIMINARY; PRT: 1407 AA.
Q8T9W5:
AC 01-JUN-2002 (Tremblrel. 21, Created)
DT 01-JUN-2002 (Tremblrel. 21, Last sequence update)
DT 01-JUN-2002 (Tremblrel. 21, Last annotation update)
DE ABC transporter Abcb2.
GN ABCB2.
OS Dictyostellium discoideum (Slime mold).
OC Eukaryota; Mycetozoa; Dictyostellida; Dictyostellium.
NCBI_TaxID=44689;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=AX4.
RA Anjaard C., Loomis W.F.
RT "Evolution of the ABC transporters of Dictyostellium."
RL Submitted (JAN-2002) to the EMBL/Genbank/DBJ databases.
DR EMBL: AF466305; NAL74249.1;
SQ SEQUENCE 1407 AA: 154763 MW: B1B6DEBEA331DAB CRC64:
Query Match 44.0%; Score 2852.5; DB 5; Length 1407;
Best Local Similarity 44.4%; Pred. No. 8.4e-157;
Matches 589; Conservative 250; Mismatches 391; Indels 97; Gaps 14;
20 KSKKNEKKKK--PVTSTFAMFRYSNMIDRLYMGTAIIHGALPLMLVFGWMTD 77
107 KKDEGEKKEGEVGPQVPFSLFRFAKPDILMITIGIALANGVSMPAISIVFGRLMN 166
78 SFANAGISRNKTPPVIIINESTNNTQHFINHEEMTYAYYYSGIAGVLYAAYIYVSF 137
167 SESPENLA-DNFDLV--EIVTSNMFYI-----ICGVFVCSYVEVAF 208
138 WCLAGROILKIRKOFHAIHROEIGMVDVHDELNTRLTLDVSKINIGIDKIGMFFH 197
209 WMLAGEQAVRCRAVYKAILKOEIIGWYDVTKSELSTRISPTLLFOBAIEKIGNFLH 268
198 STATFPTPIYGFTRG-----WKLTVLIAISPVIGLSAIIWAKILISPTDELL 247
269 HTSTLCGIFGVFGVNGKYNYYLFCQGLTVLIFALPLTAAAGAMTKMADLTKKGD 328
248 AVAKAGAAVEEVLAIKRTVIAFGCGKLEERYKNLEAKIGIKRAITANISGAAPFL 307
329 AVAKAGAAVEEKGISIRVVSFGPEVKKYTRRLLEADIDGTGKIMNGIGIGLVFLV 388
308 IYASVLAFWYGTSLVLSSEYSI-----GOVLYTFPSVLIGAFSISGQSPSITAFAN 359
389 LRFQYSLSFWYSGKLVDRKKNPVDRWQGDVLYTFEFSVIMGAMALGQASPNVASFPAN 448
360 ARGAAYEIRKIIDNKPSSIDYSKSGHK--PDNIKGNLEKKNHFSYSRKEVYKILKGLN 418

```

Db 449 GRGAFKIVYVDNRNKSIDPESTBGRSLIEFVQGNIEYRIGTSYSPRDVAFIANNMUT 508
QY 419 VOSGQYALVYNGSGCKSTTVOLMORYDPPDGMVCDIGODIFTINVRHLRETTGVSOE 478
Db 509 IKKQTVLALVDGSGGKSSVIGLIERFDPDGEVYLDGNINIKHISLRNIGLVSOE 568
QY 479 PVLEFATIAENIRYGRNTYTMDEIKAVKANAYDFIMKLPNKFDPLVGERGAOLSGGOK 538
Db 569 PVLPANIAENIRYGRNENATMDQITLACKTANADFTSALPEGYDIOVEGKGVOMSGGOK 628
QY 539 QRITAIYALVRNPKILLDEATSDLTSEAVNOVALDKARKGRTTIVIAHRLSTVRNAD 598
Db 629 QRITAIYALVRNPKILLDEATSDLTSEAVNOVALDKARKGRTTIVIAHRLSTVODAD 688
QY 599 VIAGFDGVYVEKGNHDELMKRGKGYEKLVTMOTRGNIELENAVTESESDALEMSGPK 658
Db 689 QIAVVGGAIVETGHELYALNGVYQLVNRQKGD-----DDKKKKKKKSSKSD 742
QY 659 DS-----GSSLIKRRSTRSRSHAP-----OGDRKLTGEDLNENVPVPSFWRLTK 704
Db 743 ESNNNIPSSISIDKSIQSIGADSLSETSTIGLVNDNNKKKKKKKKKKKKKKKKKKKKKKKK 802
QY 705 LNSTEMPYFVYVGCALINGGLQAPAFITISRTIIGITFTRDEDEYTRKONSMMFVLEVL 764
Db 803 LSRGDMHFLGLVGLTNGALMPAFVSTISELIGITQEDDTDLTRRSKNM-ALMFTLL 861
QY 765 GIISTFIFLQGFTEKAGEILTKRLRYVFRSMLRDVSVMFDDPKNTTGALTTRLANDA 824
Db 862 AVVALANFQIYCFTEFIEGKTLFMRRLRSLSESIMRDIQMFDLTENSRTGLTANLATEA 921
QY 825 AQVGAIGSRVAVITONINILGTGIIISLIGWQLTLLLAIPITAIAGVEMKLSGO 884
Db 922 TLVQGMTSQRLGLLQINITYIAGVIAFVSGMKLTLVLVACPVIGFGKTEMDFQGF 981
QY 885 ALKKKLEBAGKITEALIENTRTYVSLTRKQFEYMAQSLQVPRNSLRRAHIFGVSF 944
Db 982 SOKREKVAECQVASEAIGIRIVSFTCEKILKFRQCLQKPIOMSGFRKSNVSGLSF 1041
QY 945 SITQAMMFYSVAGCFRRGAVLVAN-----EFMN----- 972
Db 1042 GFSQCTLFYTLTYWYGRKLVDSGEMPAKESLTETCYNGEYANIGYDEATCIKSFYT 1101
QY 973 ---FDVILVFSATVFGAMAVGVSSFPADYAKAKVASAAHIMILKESPLIDSYSGHLK 1029
Db 1102 TEGSMMARVFPALINSAGVGQSNAPMDLGRKAKLAVALIFSLDRVSEIDPFENKQGT 1161
QY 1030 PNTLEGVNTRENVVYVPTRPDIPLVGLSLEVKKGOTLALVSSGCGKSTVYOLLERY 1089
Db 1162 LPERGDIERDKFSTPSPKNAVFQGNLVIPHGKRVLVGNSSGKSVISLTERPY 1221
QY 1090 DPLAGSVLIDKEIKHLNVQMLRAHLGIVSOEPIILFDCSIAMENIAYGDSRVVSHBEIMQ 1149
Db 1222 NPSGSLTIDGVNIKDLNLMNMLRGNMGVGOEPLFSGTIFENIYIGKPDATM--DEVE 1279
QY 1150 AAKEANIHFFLETLPEKYVTRVYDGGKOTLSGOKORITAIRALVROPHILLDEATSLAD 1209
Db 1280 AAKAANHTIETESLPDAHYTOLGDKFTOLSSGOKORAIARITIRPKVLLDEATSLAD 1339
QY 1210 TESERVOEALKABRGRCYIAHRLSTIONADLIYVFGNGYKKEHGHQOOLAOKGY 1269
Db 1340 TVSERVVOVALDVNSKGRISYIAHRLSTVIDADLIYVYKRGVVELGHEHETLAENRGY 1399
QY 1270 FSNMVSQ 1276
Db 1400 AELVSRQ 1406

```

```

DT 01-JUN-2002 (Tremblrel. 21, last annotation update)
DE Hypothetical 140.5 kDa protein.
GN C346.4.
OS Caenorhabditis elegans.
OC Eukaryota; Metazoa; Nematoda; Chromadorea; Rhabditida; Rhabditoidea;
OC Rhabditidae; Pelodierinae; Caenorhabditis.
OX NCBI_Taxid=6239;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=BRISTOL N2.
RX MEDLINE=9069613; PubMed=9851916;
RA None.
RT "Genome sequence of the nematode C. elegans: a platform for
RT Investigating biology. The C. elegans Sequencing Consortium."
RL Science 282:2012-2018(1998).
RN [2]
RP SEQUENCE FROM N.A.
RC STRAIN=BRISTOL N2.
RX Langston Y., Rohlfing T.;
RA "The sequence of C. elegans cosmid C346."
RT Submitted (APR-1997) to the EMBL/GenBank/DBJ databases.
RN [3]
RP SEQUENCE FROM N.A.
RC STRAIN=BRISTOL N2.
RX Waterston R.;
RA "Direct Submission."
RT Submitted (OCT-2001) to the EMBL/GenBank/DBJ databases.
CC -1- SIMILARITY: BELONGS TO THE ABC TRANSPORTER FAMILY.
DR EMBL: U97407; AAB52482.2; -.
DR HSSP: P13569; INED.
DR InterPro: IPR003593; AAA_Atpase.
DR InterPro: IPR001140; AbcTransportM.
DR InterPro: IPR003439; AbcTransportR.
DR Pfam: PF00664; Abc_membrane.3.
DR Pfam: PF00005; Abc_tran.3.
DR ProDom: PD000006; Abc_transportr.2.
DR SMART: SM00382; AAA.3.
DR PROSITE: PS00211; Abc_TRANSPORTER; 3.
KW ATP-binding; Hypothetical protein; Transport.
SQ
SEQUENCE 1265 AA; 140465 MW; 4948EF5CA402757 CRC64;

Query Match 45.6%; Score 2827; DB 5; Length 1265;
Best Local Similarity 45.4%; Pred. No. 2,2e-155;
Matches 583; Conservative 246; Mismatches 400; Indels 56; Gaps 14;

QY 21 KSKKNK-----KEKKPYSTFAMFRYSNMLRLVLTGTMALIHG 62
Db 2 KSKRNEPTWTKRLKRSNDSSTIDESTVAKLTNGITTYGVDLILITGVAAVHG 61
QY 63 AALPLMLVFGNNYDSFANAGISRNKTPPV-IINESITNNQHFIMHLEEMTYAAYYS 121
Db 62 AGFPLLAIVGGMTVFELRA--QNSDFVGVVDVNPBGLVPLSIDFENSEVVKCYIYL 118
QY 122 GIGAGVVAANYIOVSPWCLAGQOLIKTRQFPAHMRQIEGFWFDVHDVGEINTRLTDDV 181
Db 119 VLVGLMFTSYVGLAEEFSYAEVLVHKLRQNYLKAIIIRQDIQGFQDQGNLTARLTDL 178
QY 182 SKIMEGIGDKIGFHSITATFTFTGTVGFTGKRLTYIATISPVGLSAAIMATLSSF 241
Db 179 ERVREGIGDFALLVOMFAALAGVGFYSWMTLVMMGFAPLLVLVLSGAKMSSMSTR 238
QY 242 TDKELLAVNAGAVAEVLAIRTYIAGGOKLEERYNNLEBEAKGIGIKRAITANISI 301
Db 239 TRVBOETIYAVAGAIAREFTSSIRYVHSLNGHKRLDRFYNALFVGQOTIVKYCYMGIGV 298
QY 302 GAAPLLIYASVIALAFWYGSVLSS--EKSIGQVLTVEFSVLIGAFSIGQASISAPANA 360
Db 299 GFSNLCMYSYALAFWYGSTLIINDPTFDRGLITTYFAVALSSGTSIGGALRHLSFGTA 358
QY 361 RGAAYEIKTIIDKPSIDYSKSGHKPDNIGKLEFNANHFSTPSKREKYLKIGLNLKVQ 420
Db 359 RGAASYLRVINSHPKIDYSLLEGILVNMKGDISFDVHFYPSRKDIHVLGISTLELK 418

```

Tue Dec 10 07:55:22 2002

us-09-672-725c-2.rspt

Page 14

Qy	421	SGOTVALVNSCCGKSTTVVOLMROKVDLPDQWAVIC-IDGODITIVRILRETTGVSOEY	480
Db	419	AGCKIALVSSCCGKSTTVNLDRFDPRDKRVLILDGVDLRENVVHSLRQIGVISOEY	478
Qy	481	LEPTTIAEMIRIRGRNNTYMDIEKAKVANAADFTMKPKNEPDLVYREBQAOLSGOKOR	540
Db	479	LEPGTTEYIKKNGNHATHHDOVYACKAMNADFTKRLPDLGIRGVEKQVJOLSGOKOR	538
Qy	541	IATARLVNRPKILLDEATSDLTSEAVNOVALDKARKGRTTVIAHRLSTVRNADYI	600
Db	539	IATARLVNRPKILLDEATSDLTSEAREVOGALDOAQARTTVIAHRLSTVRNADRI	598
Qy	601	AGFDODVYIEKENHDELMEKREKGYFKLV-----TWOTRNETELENAIGSEKSEDAL	653
Db	599	FYVKAGNIYESSGHELMKSGKGFYDMTOAQVVRQOQOEAKDIE-----DTISMSAHS	652
Qy	654	EMBRKSSGSLKRRSTRSTRSHAPQODRKLGTREKDLNE-NVPPYSHWRLKLNSTEMPY	712
Db	653	HLSRKSTSTSAI---STATSHO-----LAEEVEECKAPTSMSEKTFEKNQDKGW	700
Qy	713	FVVGIGIATINGLOPAPSLIFSNIGITFREDDEPTRKONSNFVSLFVLGILISFTE	772
Db	701	FIGIGIGAFIESFVYPAVALYAEIFNVYSLPAD---QOQAVVYFMCMEFLMTGITEFYGF	758
Qy	773	FLOGFPGKAGGIIILKRLRYVPSFSLMDODVEMPDOPKNTGATLTTPDLANDAAQVKAIG	832
Db	759	FTSANGLGRGSGELMKRFEARFKNLLRQDIAPFYDRLHNGIKGLCTRFATADAPNR-VYF	817
Qy	833	SRILAVTQIANLGGIILISLIGYMOZTLILAIYPIAINGVEMKLSGALDKDEL	892
Db	818	TRLPVVLASTVITGALIGIYGYOMQALILVWAPLVLVMOGYIEBMQMRGKIRTOQL	877
Qy	893	EGAGKATATEALINFTVVSILTRDEQFEYMTAOSLOVPRYNSLRKAHIEGVESFTIOAMXY	952
Db	878	EEAGKVASQAVHEITVYSILNRDOPHFYCEYLEPENTLILKAHYGVAFESQSILF	937
Qy	953	FSTVAGCFRGALVANEHMFNOVDLVLVFSAYIFGMAAGVYSSFPADYAKVSAHAHYIM	1012
Db	938	FMYAAAFYLGISLFVNOQAMOPIDVYREFRAISFCQMGIMGNTFSILPDUVAKRLASLLEY	997
Qy	1013	TIKSPDLIDSYSPHFL-KPNTLEBQNTYFNVGFYENPYPBPDLPVLOGSLSEVKKGGOTLAV	1071
Db	998	LIEHPPIIDLSDSOIVPIT--GNSISINVFENFPHKQIKVLOGGTIDIRAKGKVALY	1055
Qy	1072	GSSGCGKSTVOLLERFYDPLAGSVLIDGKEIKHLNVQOMLPAHLGIVSOEPIIFDCSTAE	1131
Db	1056	GHSGCCGKSTYMWLERFYNODKGMIMDINRINISLSLEQVCIVSOPITPDTGGE	1115
Qy	1132	NIAYGNSNVSVSHEIMQAKKANIHHTETLEPEKYNNRVQDKTQOLSGOKORIALARA	1191
Db	1116	NICYGTN-RNMYVCEIYEAKKAMNINHEILGPDGYDHYHGEKQOLSGOKORIALARA	1174
Qy	1192	LVNROPHILLDATSDLTSEKVOYOBALDKAREERTQIVIAHLSLTIONADLIYVPONG	1251
Db	1175	LVNPSVLLLDATSDLTSEKIVQOBALDAKQGRCTLVIAHLSLTIONSDVIALYVSG	1234
Qy	1252	KYKEHGTHOOLLAQKGIYFSMWVQ	1276
Db	1235	KIVEKGTDBELIRKSEIYOKFCEYO	1259

Search completed: December 9, 2002, 16:58:44
Job time : 1323 secs